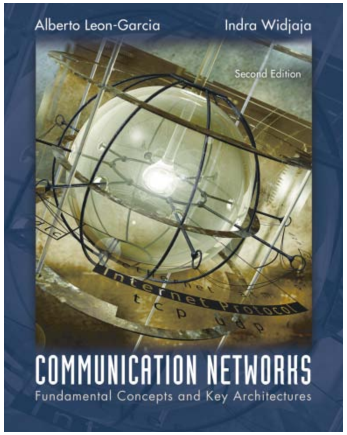
Chapter 7 Packet-Switching Networks

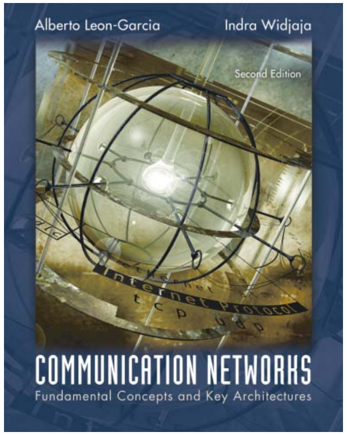


Network Services and Internal Network Operation

Packet Network Topology

Datagrams and Virtual Circuits

Chapter 7
Packet-Switching Networks

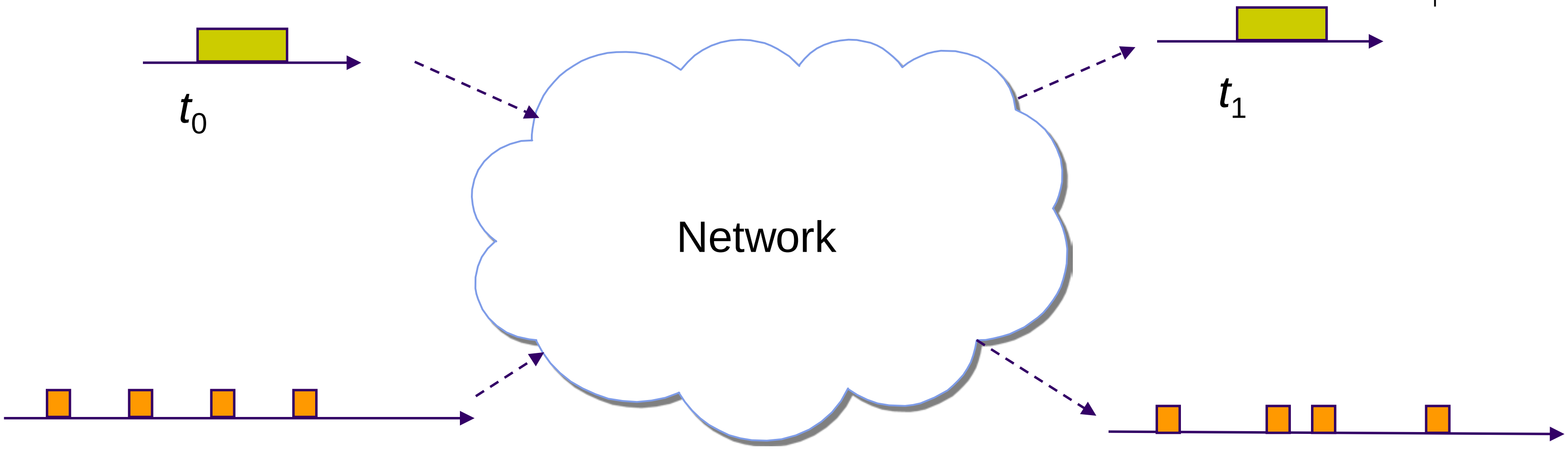


Network Services and Internal Network Operation

Network Layer

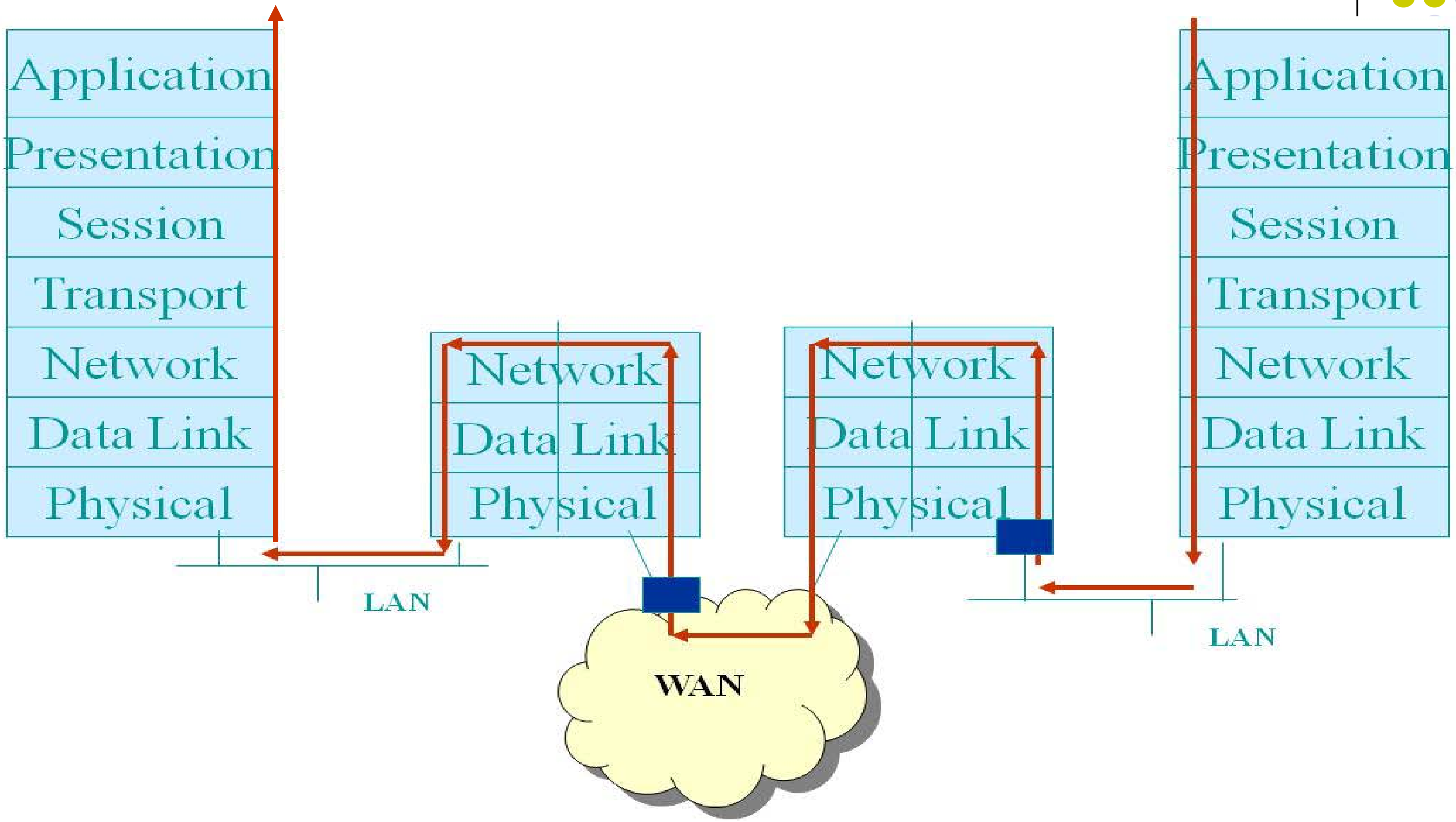
- Network Layer: the most complex layer
  - Requires the coordinated actions of multiple, geographically distributed network elements (switches & routers)
  - Must be able to deal with very large scales
    - Billions of users (people & communicating devices)
  - Biggest Challenges
    - Addressing: where should information be directed to?
    - Routing: what path should be used to get information there?
    - Efficiency: how to forward the sheer volume of traffic?

Packet Switching



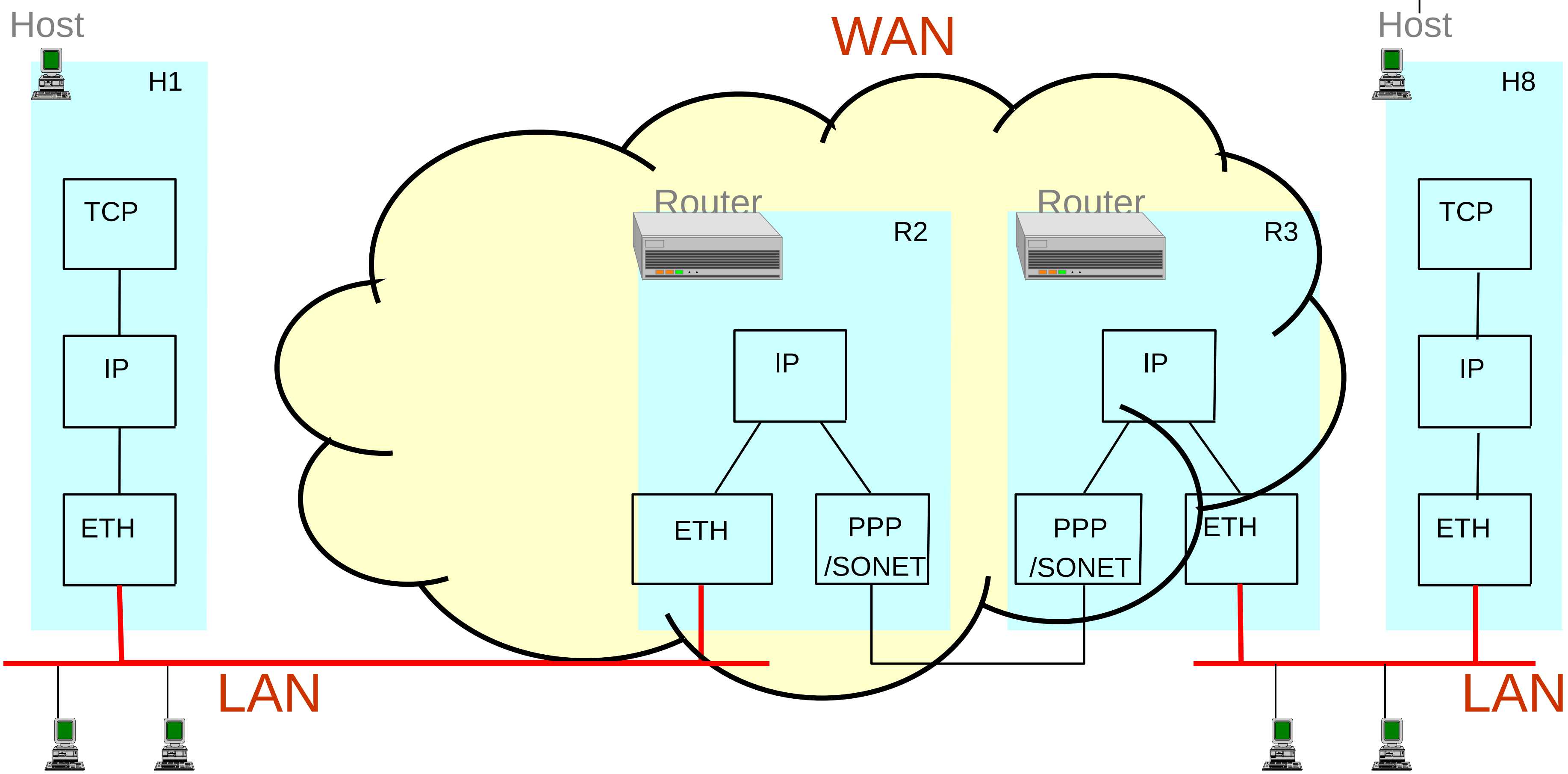
- Transfer of information as payload in data packets
- Packets undergo random delays & possible loss
- Different applications impose differing requirements on the transfer of information

Network Service

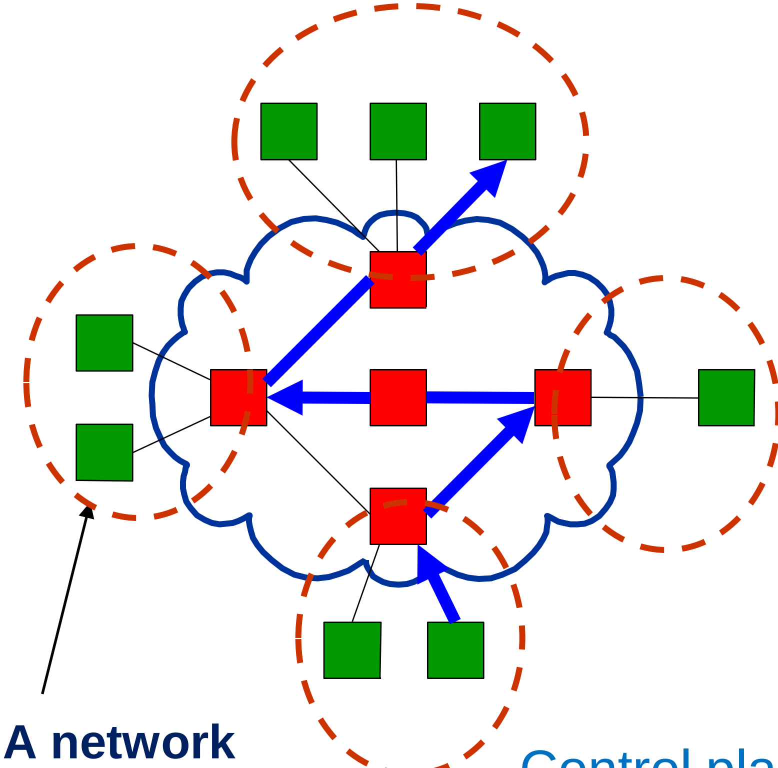


- Network layer can offer a variety of services to transport layer
- Connection-oriented service or connectionless service
- Best-effort or delay/loss guarantees

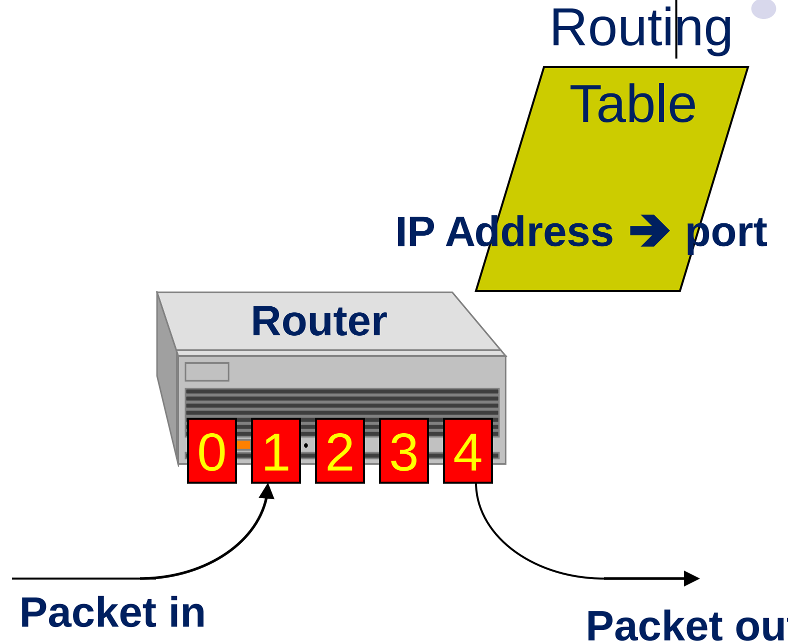
Interworking/Internetworking



IP = Network Layer



A network



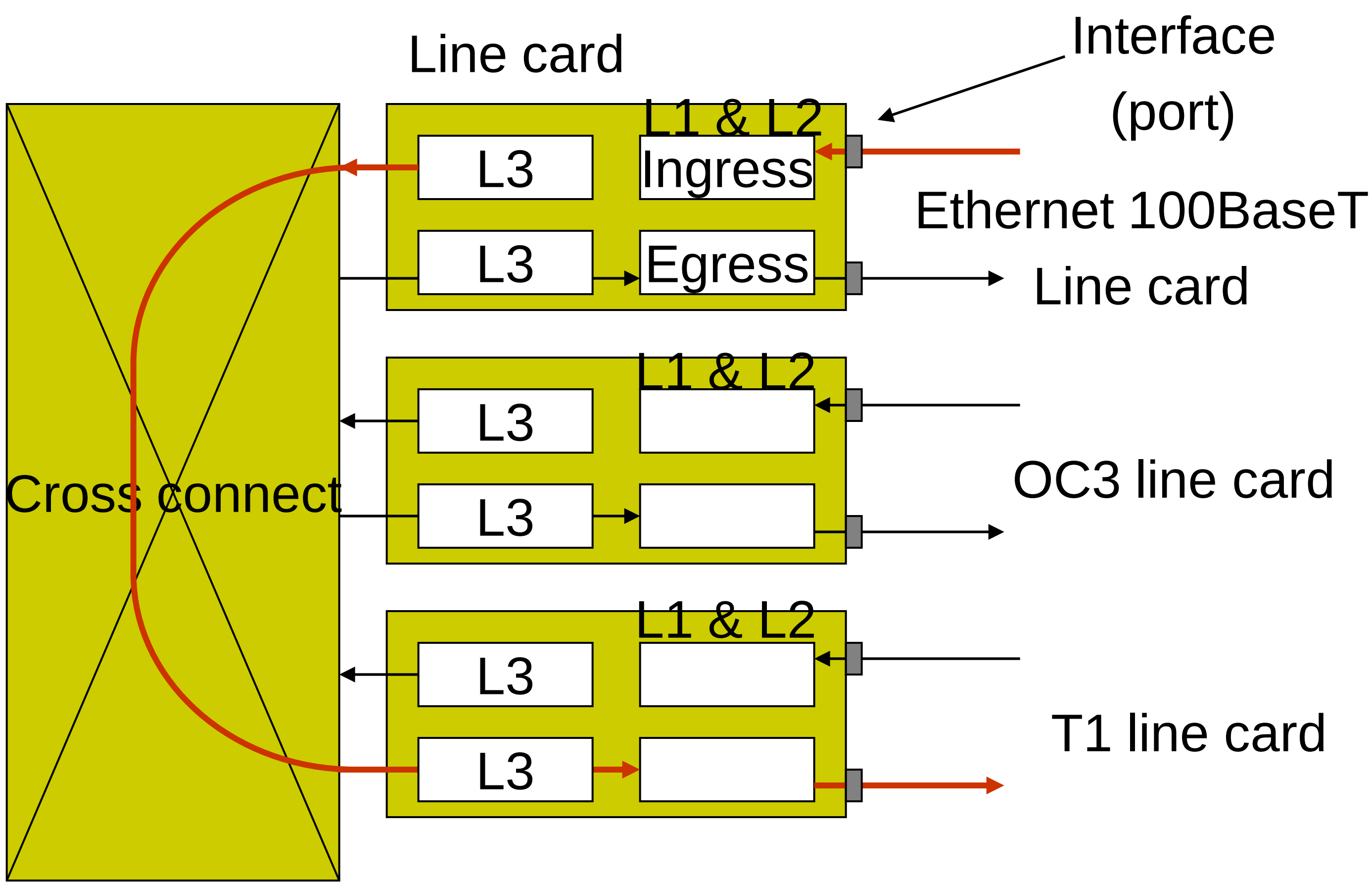
Control plan vs. data plan:

1. How to build routing table?
2. What information to carry in the packet header?
3. How to use this information together with Routing table to forward packets?

Router

Host

Router Generic Architecture



Network Service vs. Operation

Network Service

- Connectionless
  - Datagram Transfer
- Connection-Oriented
  - Reliable and possibly constant bit rate transfer

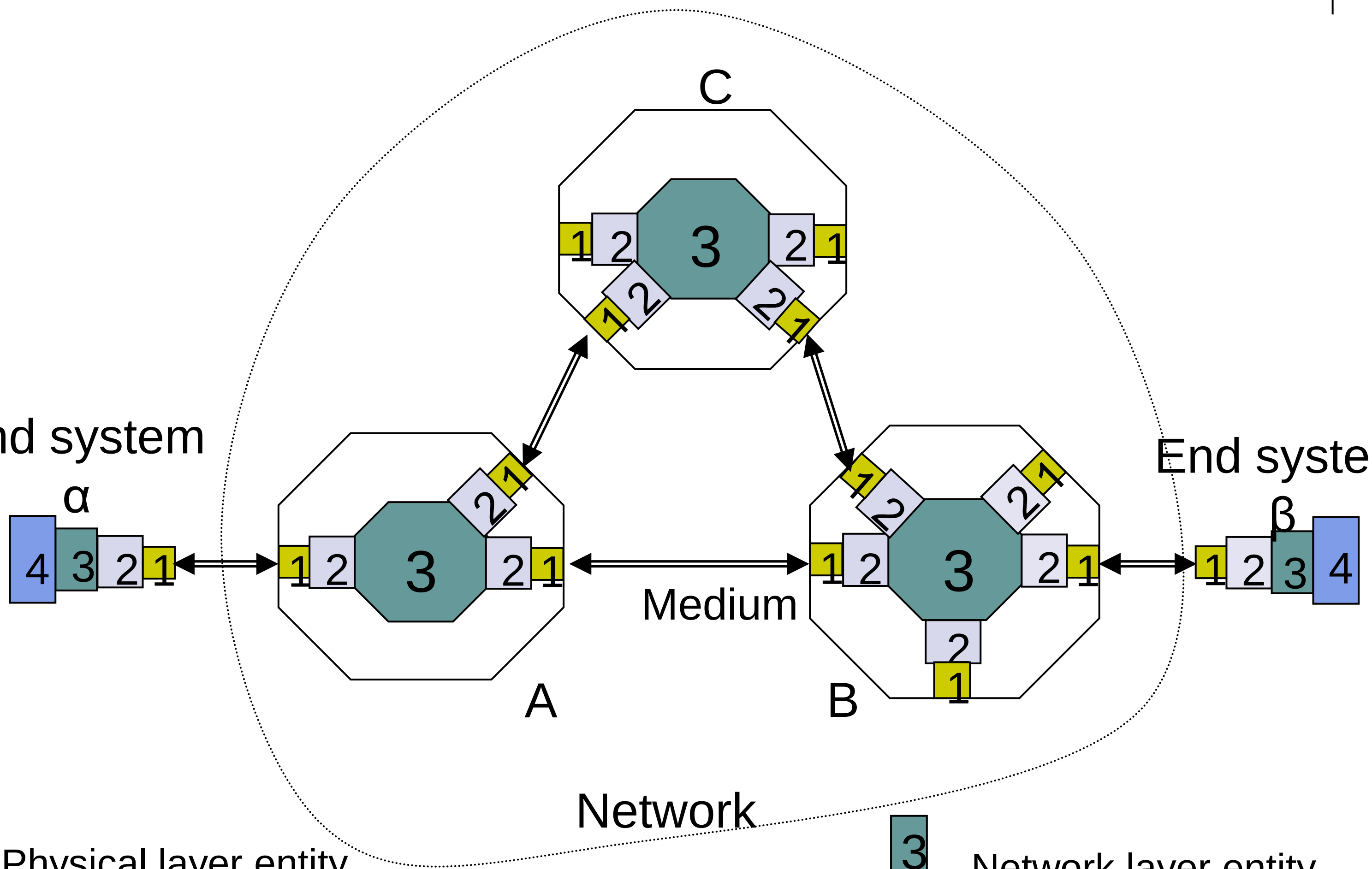
Network Operation

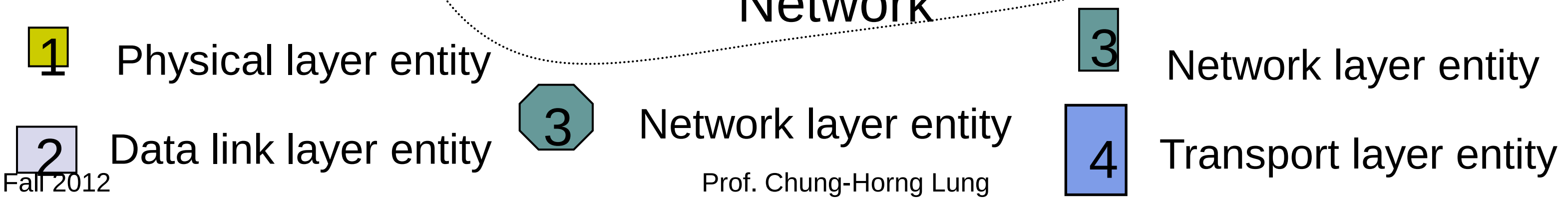
- Connectionless
  - IP
- Connection-Oriented
  - Telephone connection
  - ATM

Various combinations are possible

- Connection-oriented service over Connectionless operation
- Connectionless service over Connection-Oriented operation
- Context & requirements determine what makes sense

Complexity at the Edge or in the Core?





The End-to-End Argument for System Design

- An end-to-end function is best implemented at a higher level than at a lower level
  - End-to-end service requires all intermediate components to work properly
  - Higher-level better positioned to ensure correct operation
- Example: stream transfer service
  - Establishing an explicit connection for each stream across network requires all network elements (NEs) to be aware of connection; All NEs have to be involved in re-establishment of connections in case of network fault
  - In connectionless network operation, NEs do not deal with each explicit connection and hence are much simpler in design

Network Layer Functions

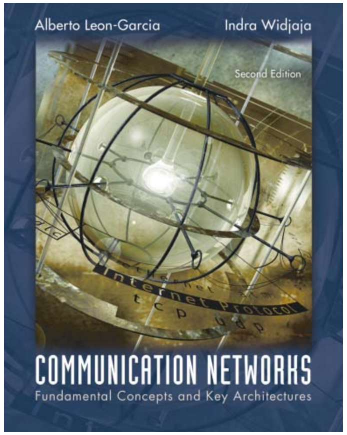
Essential:

- **Routing**: mechanisms for determining the set of best paths for routing packets requires the collaboration of network elements
- **Forwarding**: transfer of packets from NE inputs to outputs
- **Priority & Scheduling**: determining order of packet transmission in each NE

Others (examples):

- Signaling, traffic engineering
- Protection and restoration
- Virtual private networks

Chapter 7
Packet-Switching Networks

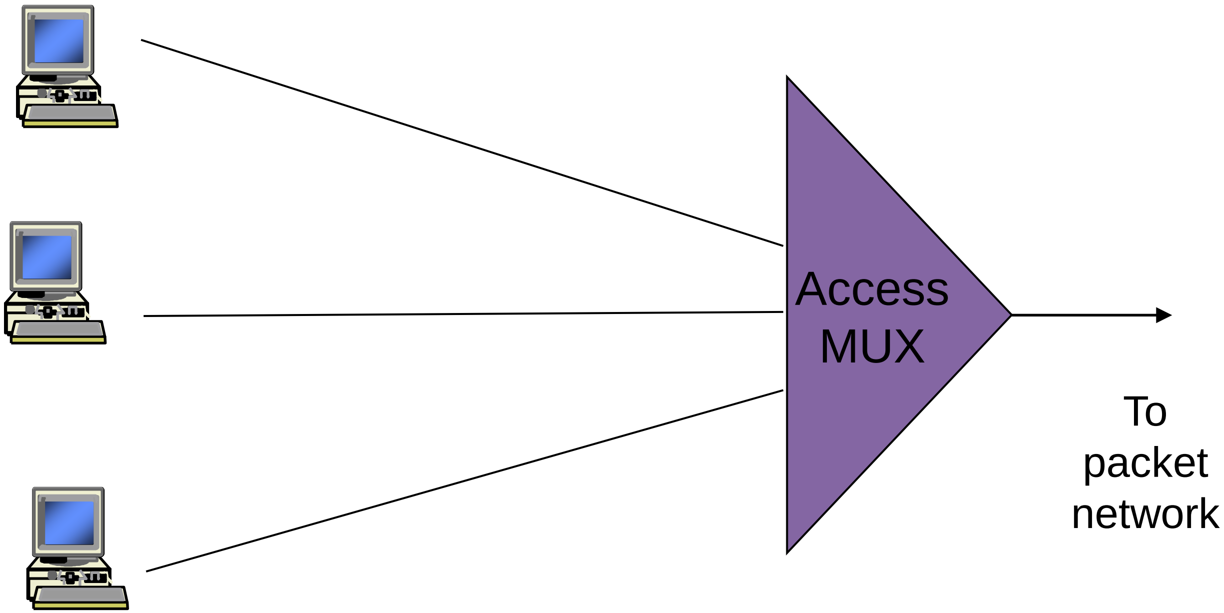


Packet Network Topology

End-to-End Packet Network

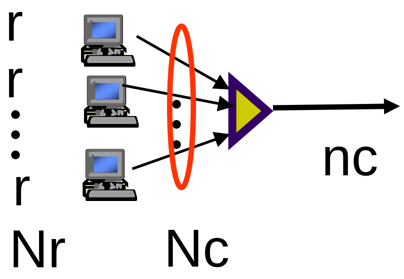
- Packet networks (packet switching) very different than conventional telephone networks (circuit switching)
- Individual packet streams are highly bursty
  - Statistical multiplexing is used to concentrate streams
- User demand can undergo dramatic change
  - Peer-to-peer applications stimulated huge growth in traffic volumes
- Internet structure highly decentralized
  - Paths traversed by packets can go through many networks controlled by different organizations
  - No single entity responsible for end-to-end service

Access Multiplexing



- Packet traffic from users multiplexed at access to network into aggregated streams
- DSL traffic multiplexed at DSL Access Mux
- Cable modem traffic multiplexed at Cable Modem Termination System

Oversubscription and Statistical Multiplexing

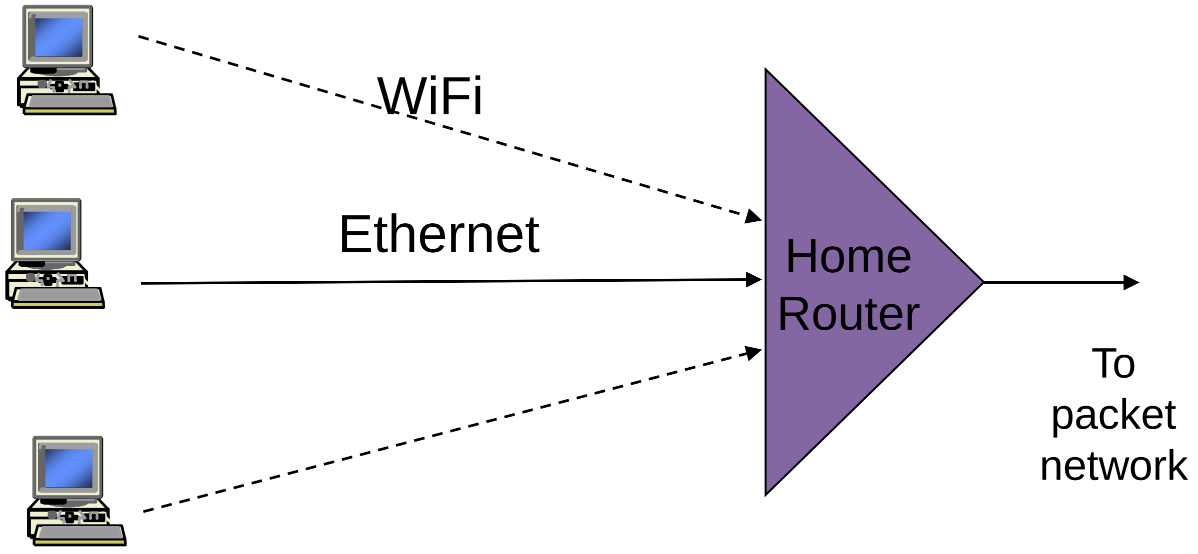


- Access Multiplexer
  - N subscribers connected @ c bps to mux
  - Each subscriber active r/c of time
  - Mux has C=nc bps to network
  - Oversubscription rate: N/n
  - Find n so that at most 1% overflow probability

Feasible oversubscription rate increases with size

N	***r/c***	***n***	***N/n***	
10	0.01	1	10	10 extremely lightly loaded users
10	0.05	3	3.3	10 very lightly loaded user
10	0.1	4	2.5	10 lightly loaded users
20	0.1	6	3.3	20 lightly loaded users
40	0.1	9	4.4	40 lightly loaded users
100	0.1	18	5.5	100 lightly loaded users

Home LANs



- Home Router
  - LAN Access using Ethernet or WiFi (IEEE 802.11)
  - Private IP addresses in Home (192.168.0.x) using Network Address Translation (NAT)
  - Single global IP address from ISP issued using Dynamic Host Configuration Protocol (DHCP)

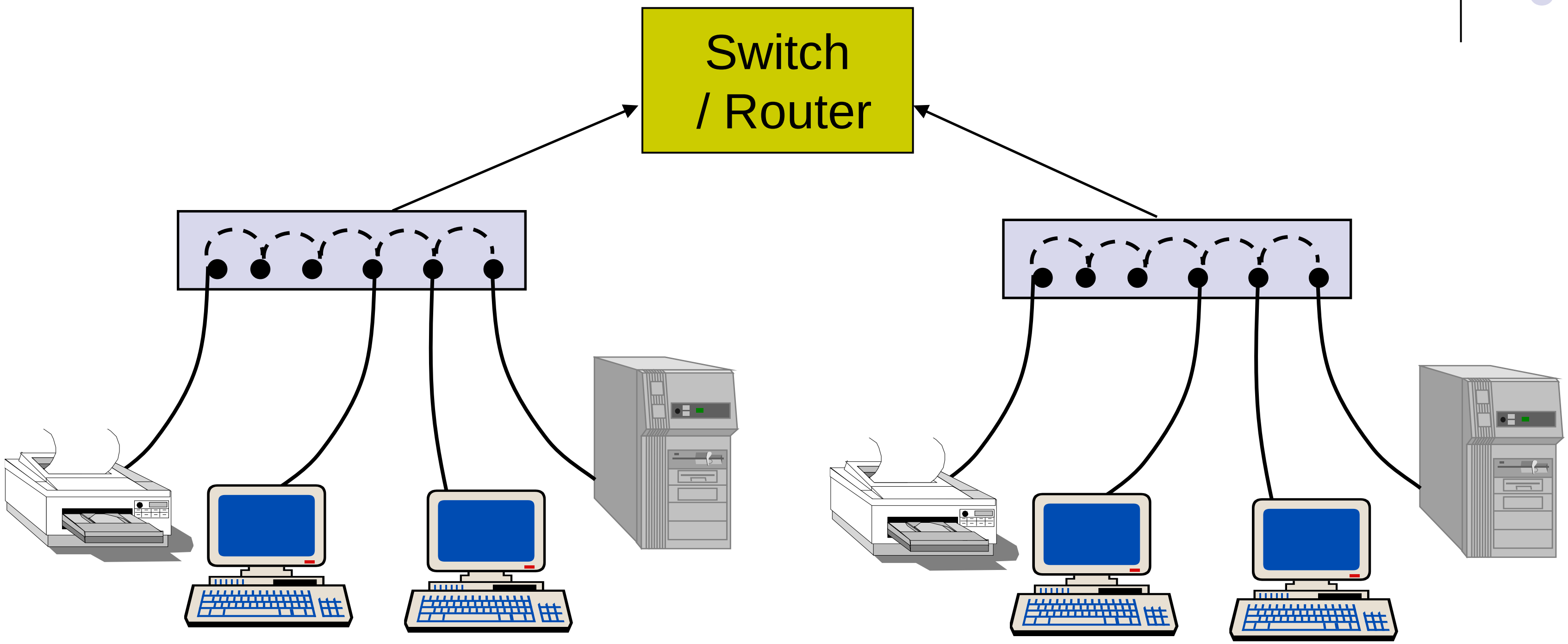
Network Address Translation

- NAT is used for ISPs to assign a single global network address to the subscriber in order to conserve address space.
- NAT converts a private address (only defined in a home network or enterprise network) to a global network address when a packet leaves the home (enterprise) network and vice versa when a packet arrives at the home (enterprise) network.

DHCP – Dynamic Host Configuration Protocol

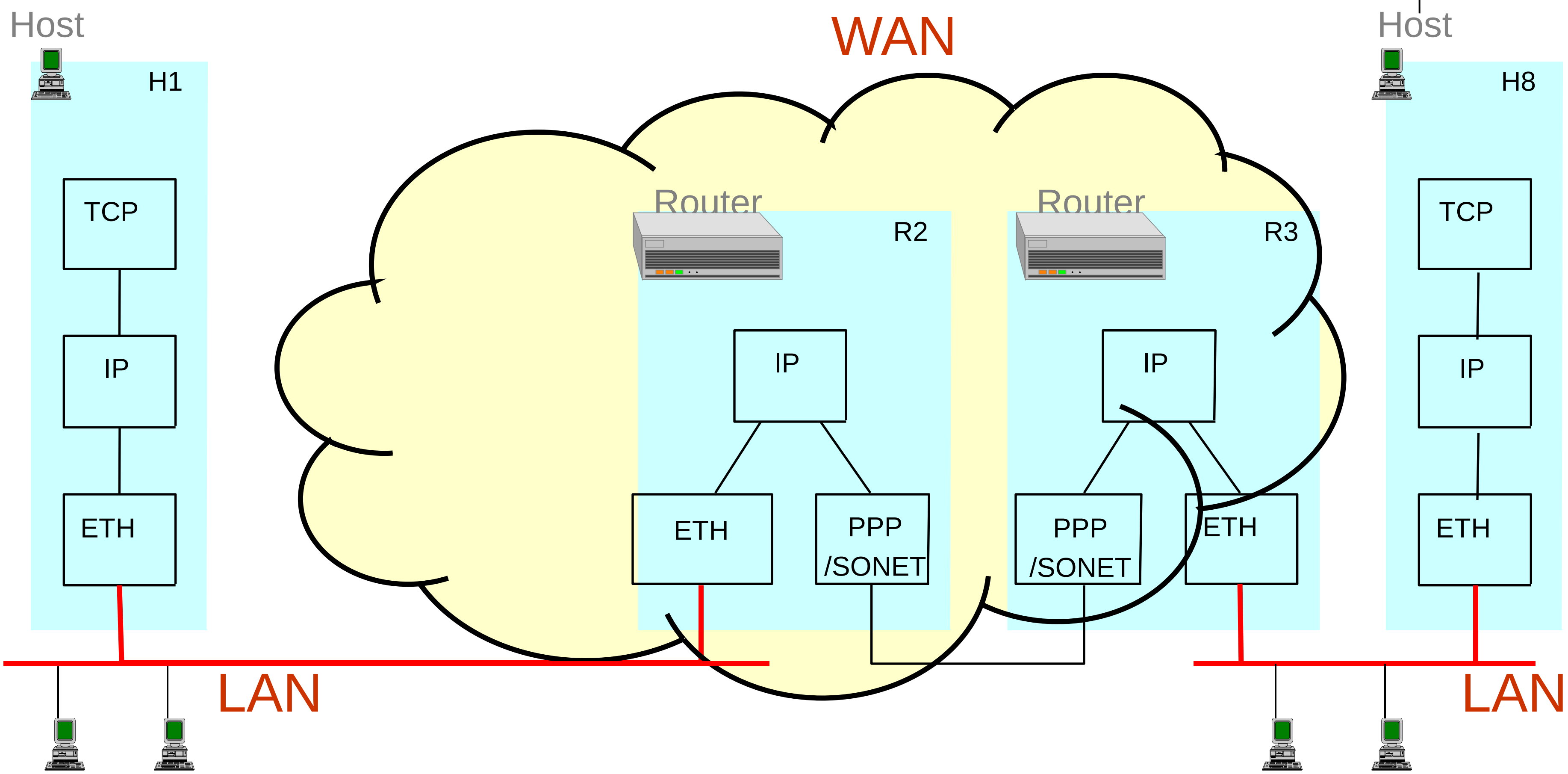
- Centralized repository of configuration data for all clients (hosts) on network
- Host Configuration options:
  - Can be manually configured (hardware address vs. configuration), but could be error prone
  - Dynamic Host Configuration Protocol (DHCP) makes life easy:
  - Host
    - DHCP server discovery “DHCPDISCOVER” broadcast message on boot up (Plug & Play)
  - DHCP server (or relay agent) responds
  - Dynamic address assignment from an address pool
    - Address reuse
    - Private address - networks 10. And 192.168
  - DHCP packet is carried over UDP

LAN Concentration

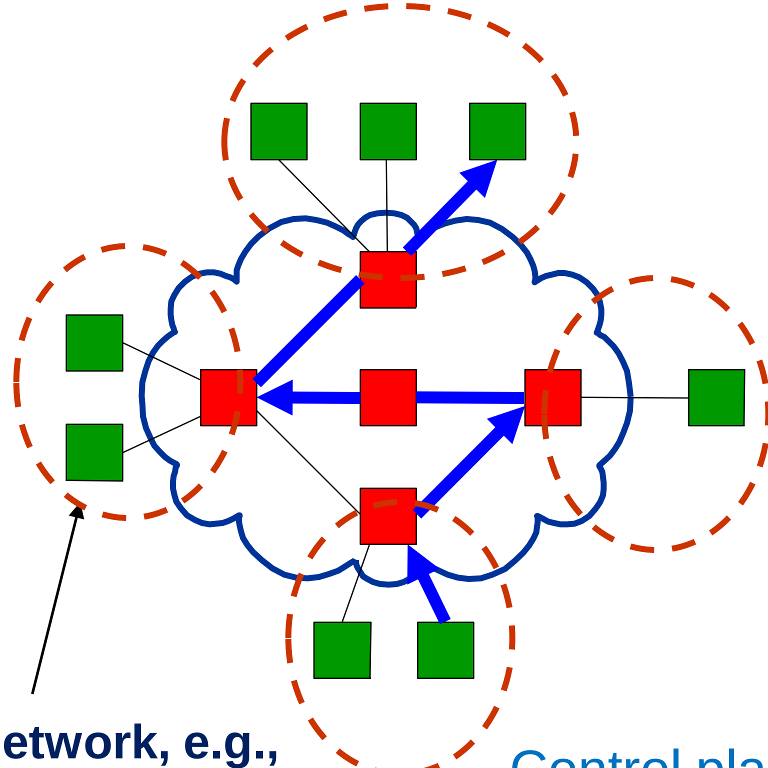


- LAN Hubs and switches in the access network also aggregate packet streams that flows into switches and routers

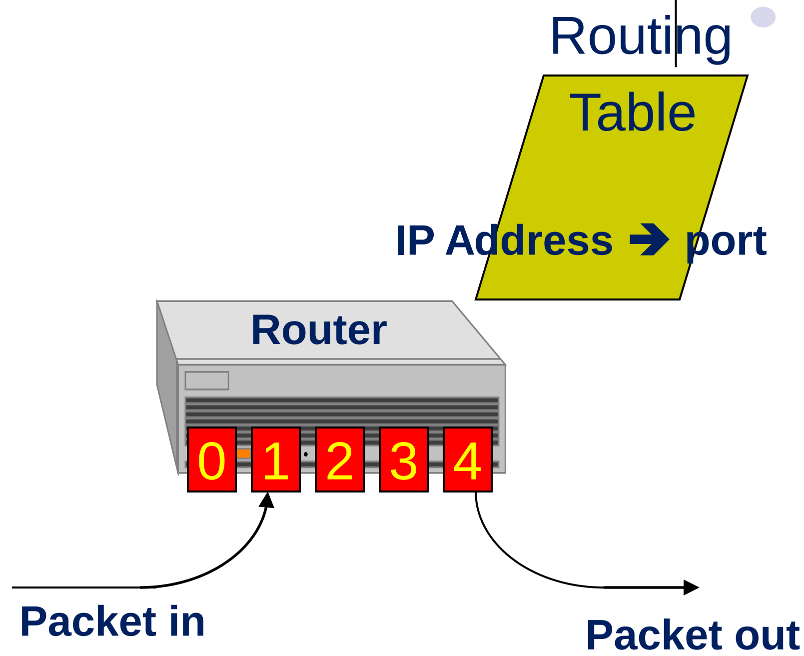
Interworking/Internetworking



IP = Network Layer



A network, e.g., campus network



Router

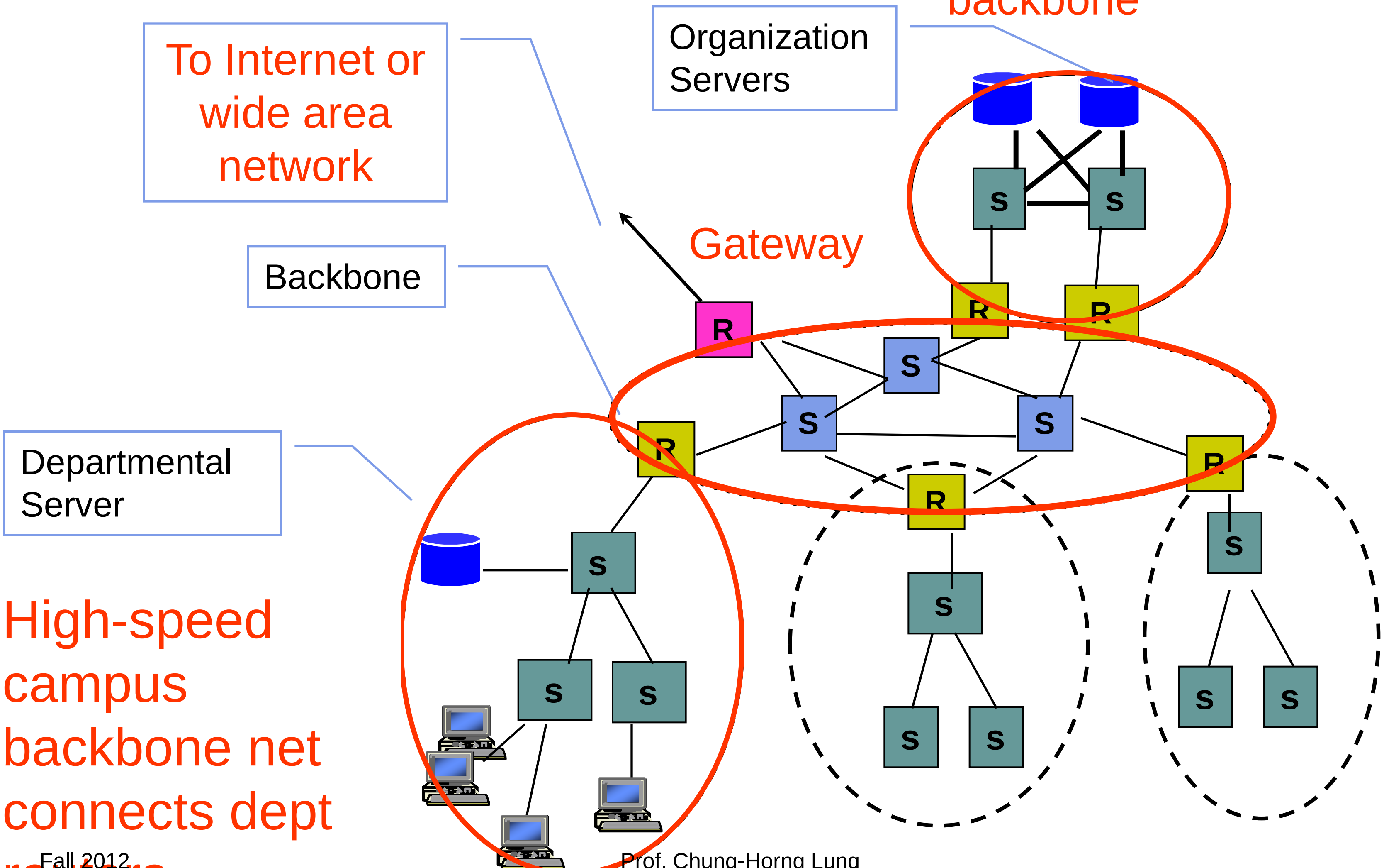
Host

Control plan vs. data plan:

1. How to build routing table?
2. What information to carry in the packet header?
3. How to use this information together with Routing table to forward packets?

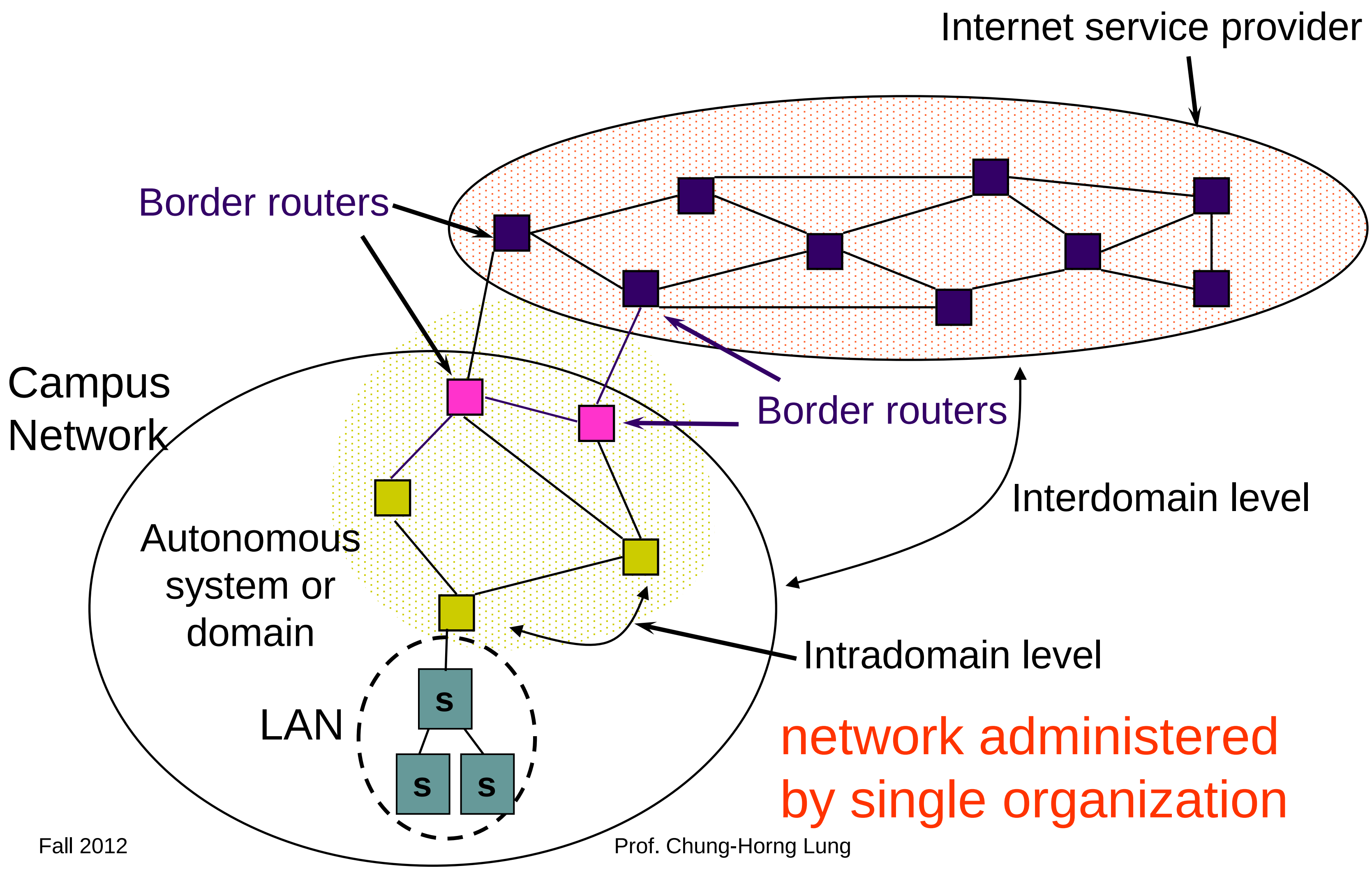
Campus Network

Servers have redundant connectivity to backbone

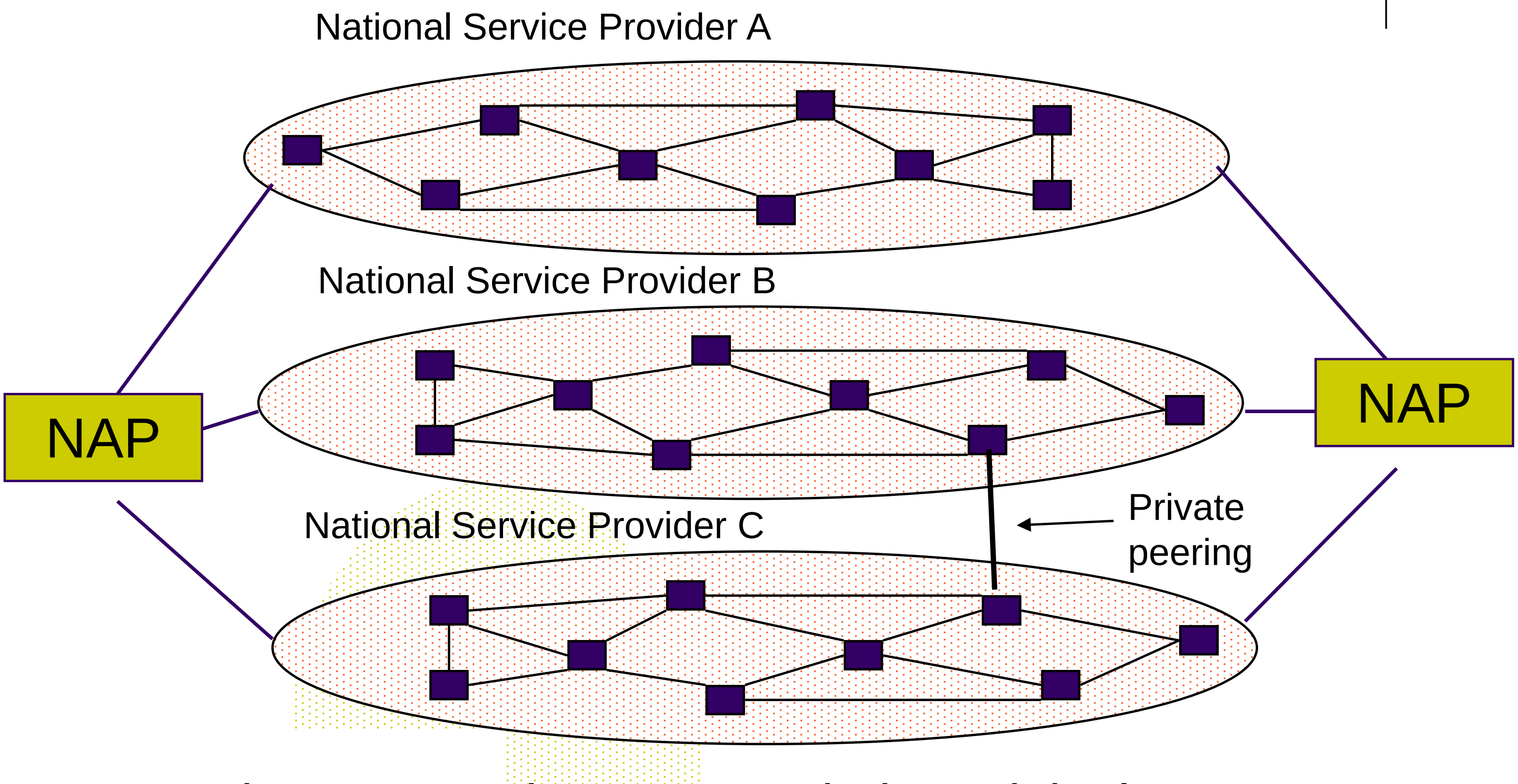


High-speed campus backbone net connects dept routers

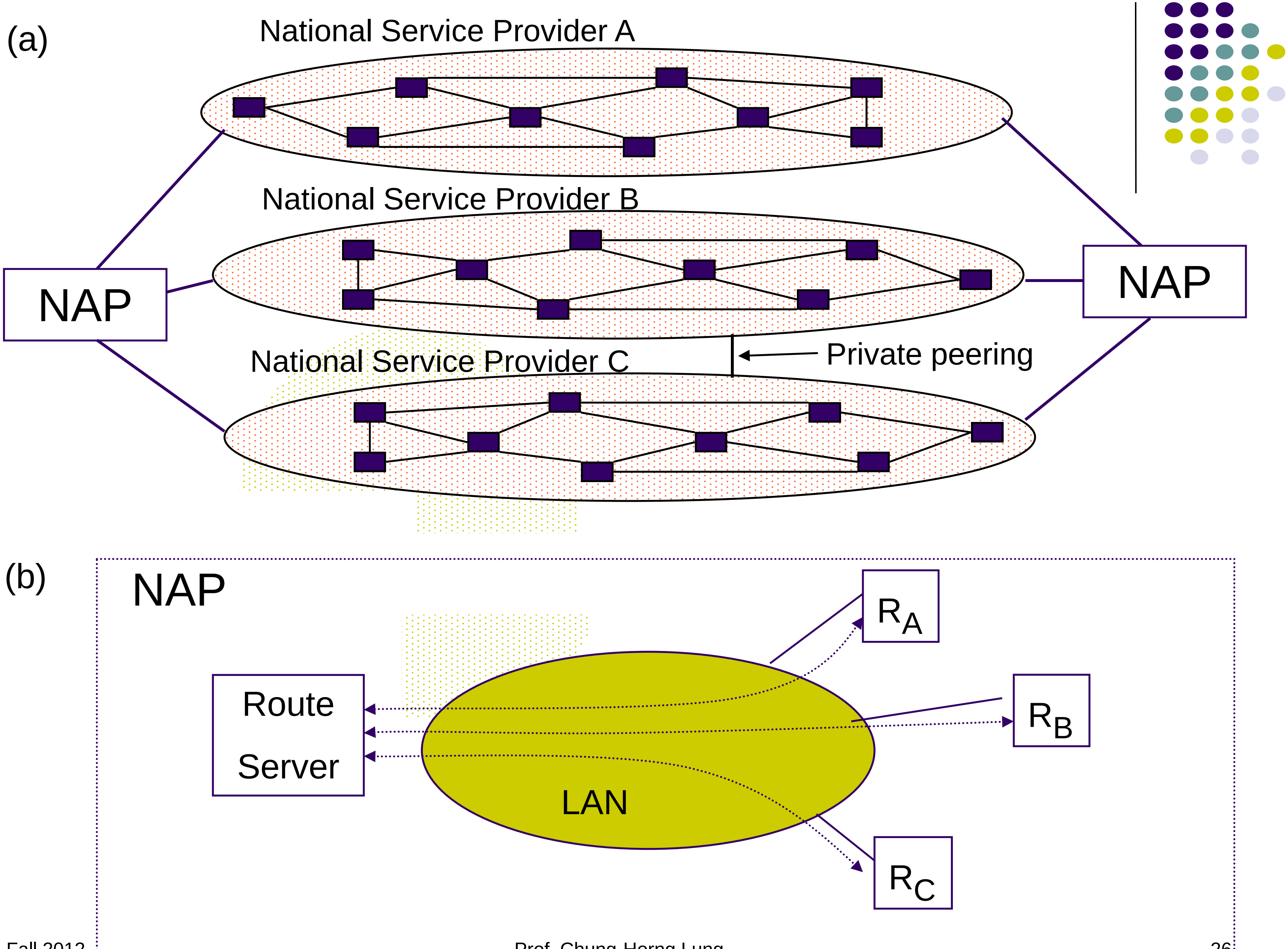
Connecting to Internet Service Provider



Internet Backbone



- Network Access Points: set up during original commercialization of Internet to facilitate exchange of traffic
- Private Peering Points: two-party inter-ISP agreements to exchange traffic

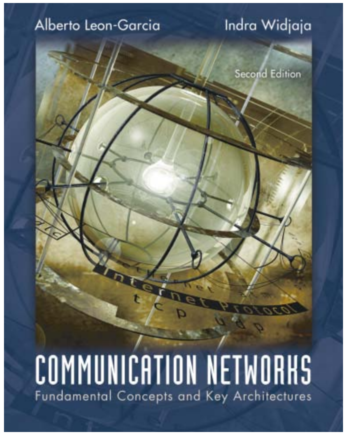

(a)
National Service Provider A
National Service Provider B
National Service Provider C
NAP
NAP
Private peering
(b)
NAP
Route
Server
LAN
R_A
R_B
R_C

Key Role of Routing

How to get packet from here to there?

- Decentralized nature of Internet makes routing a major challenge
  - Interior gateway protocols (IGPs) are used to determine routes within a domain
  - Exterior gateway protocols (EGPs) are used to determine routes across domains
  - Routes must be consistent & produce stable flows
- Scalability required to accommodate growth
  - Hierarchical structure of IP addresses essential to keeping size of routing tables manageable

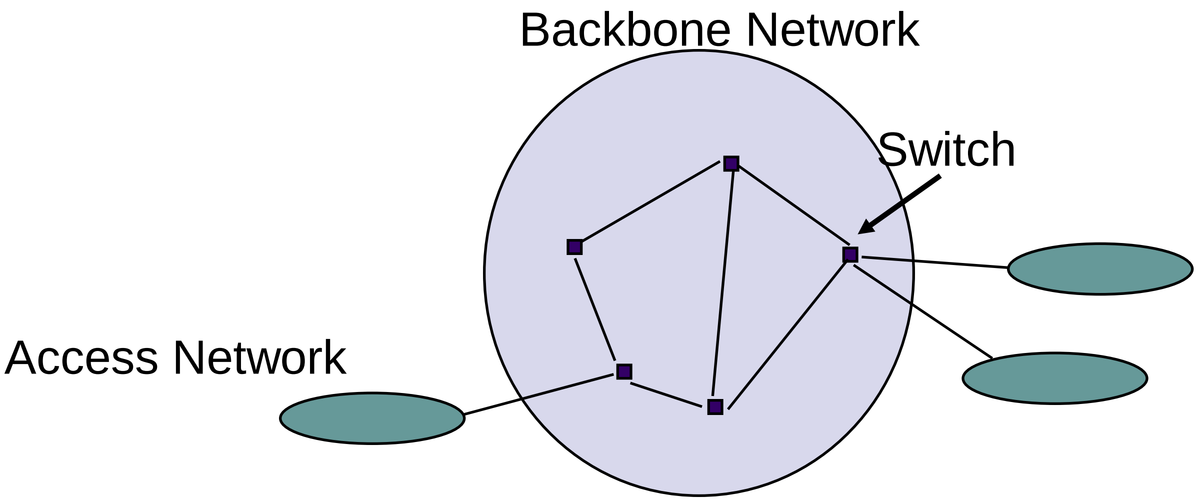
Chapter 7
Packet-Switching Networks



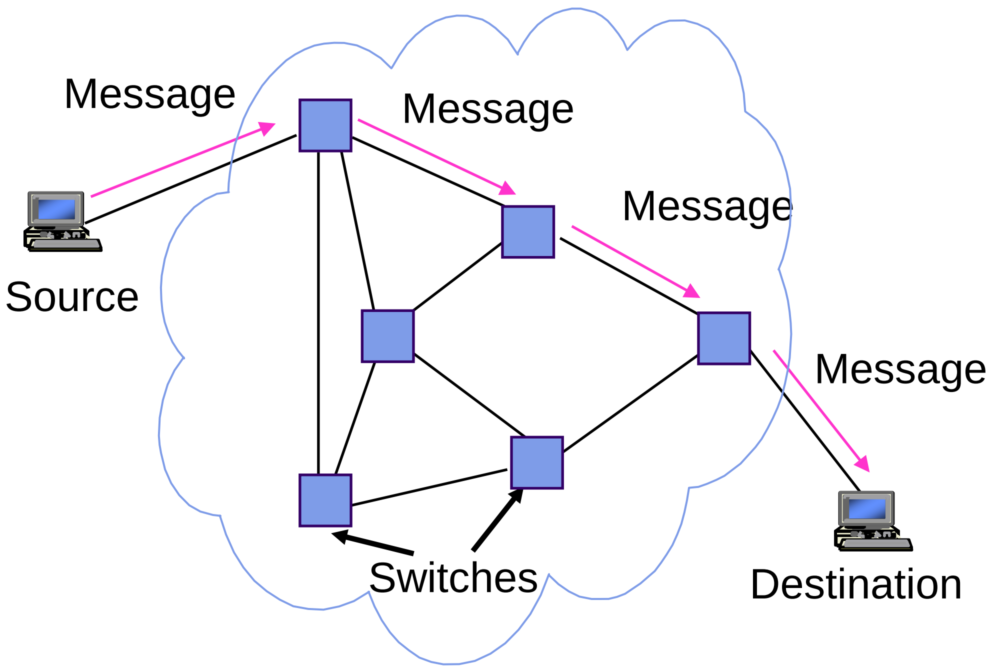
Datagrams and Virtual Circuits

The Switching Function

- Dynamic interconnection of inputs to outputs
- Enables dynamic sharing of transmission resource
- Two fundamental approaches:
  - Connectionless
  - Connection-Oriented: Call setup control, Connection control

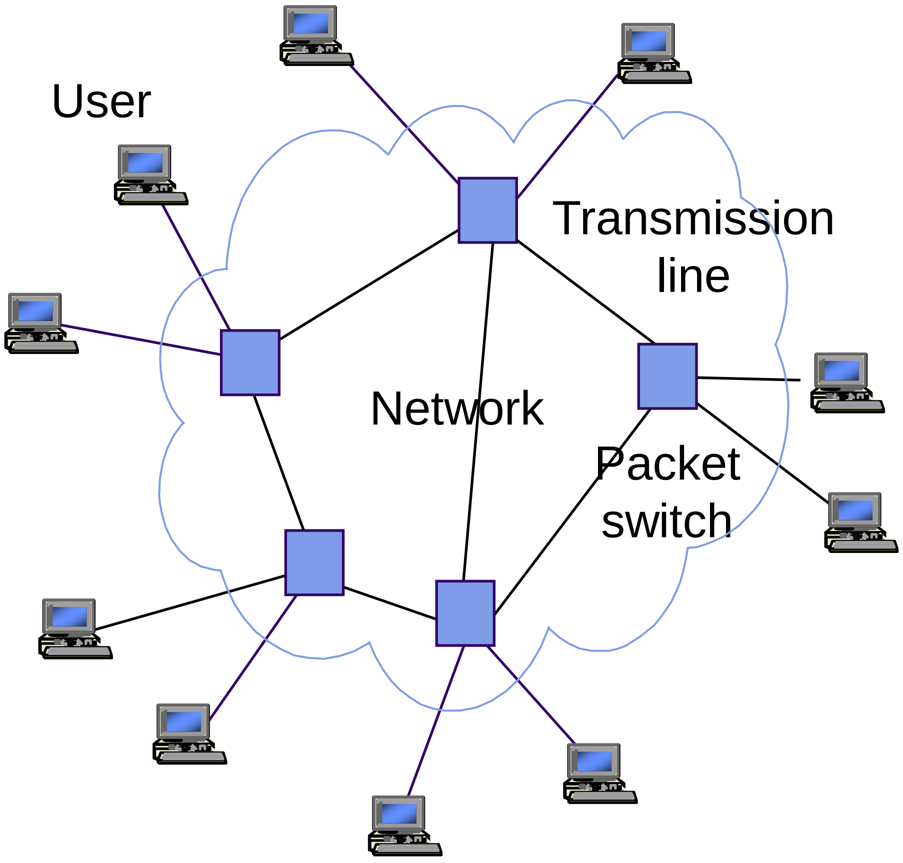


Message Switching



- Message switching invented for telegraphy
- Entire messages multiplexed onto shared lines, stored & forwarded
- Headers for source & destination addresses
- Loss of messages may occur when a switch has insufficient buffering to store the message

Packet Switching Network



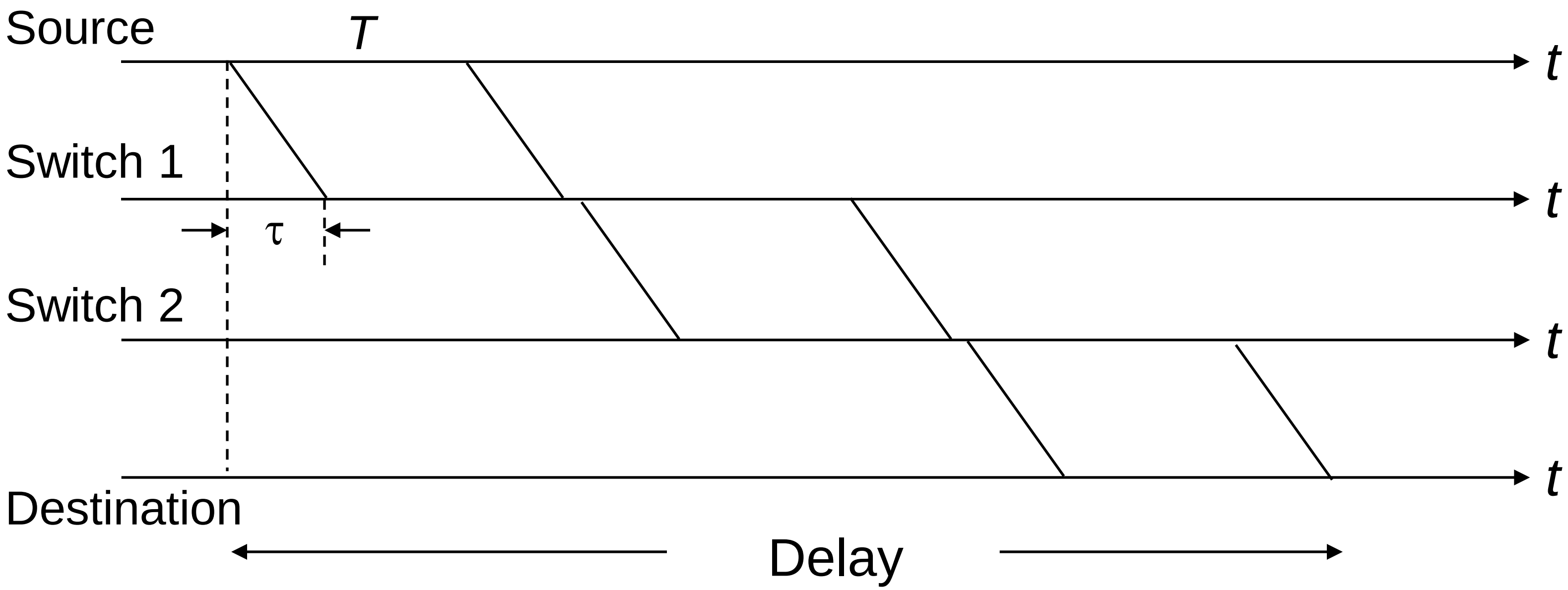
Packet switching network

- Transfers packets between users
- Transmission lines + packet switches (routers)
- Origin in message switching

Two modes of operation:

- Connectionless
- Virtual Circuit

Message Switching Delay



Minimum delay = $3\tau + 3T$

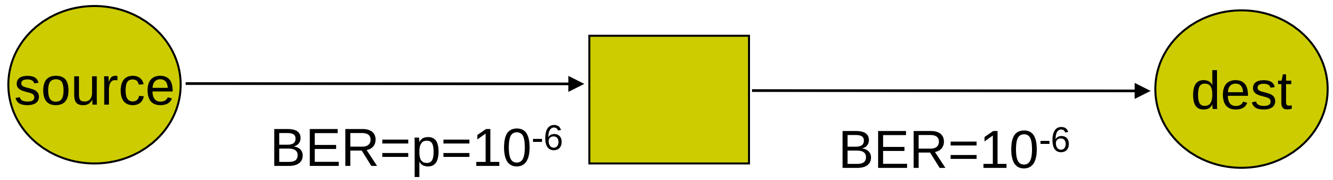
τ: propagation delay

T: transmission delay

Additional queueing delays possible at each link

Long Messages vs. Packets

1 Mbit message



How many bits need to be transmitted to deliver message?

- Approach 1: send 1 Mbit message
- Probability message arrives correctly

$$P_c = (1-10^{-6})^{10^6} \approx e^{-10^6 10^{-6}} = e^{-1} \approx 1/3$$

- On average it takes about 3 transmissions/hop
- Total # bits transmitted ≈ 6 Mbits

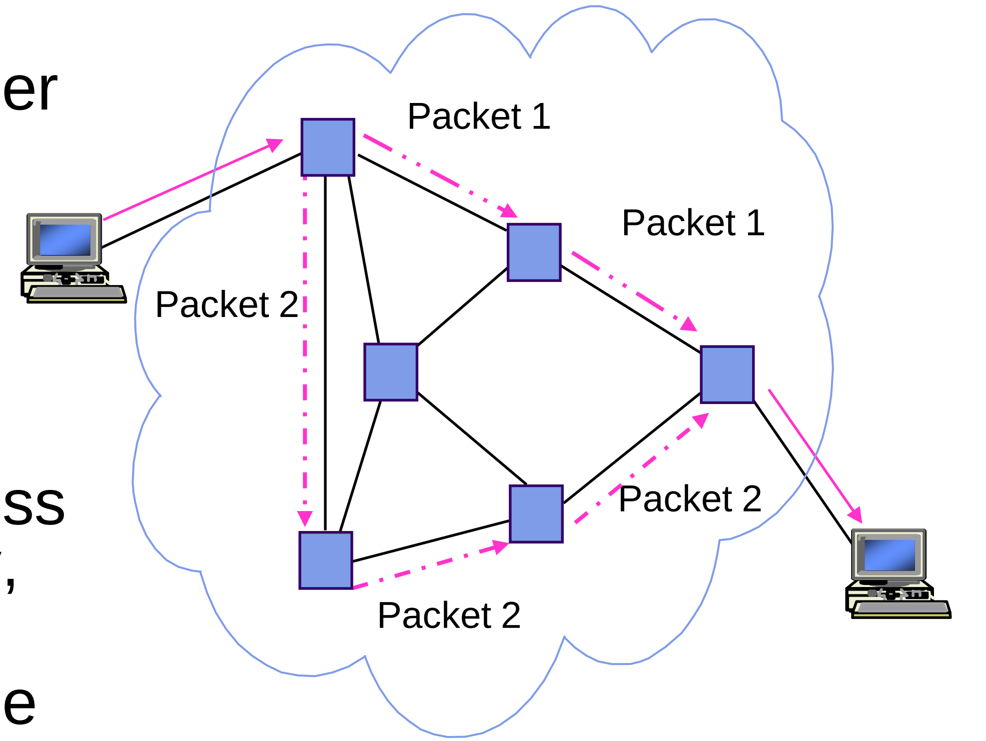
- Approach 2: send 10 100-kbit packets
- Probability packet arrives correctly

$$P_c' = (1-10^{-6})^{10^5} \approx e^{-10^5 10^{-6}} = e^{-0.1} \approx 0.9$$

- On average it takes about 1.1 transmissions/hop
- Total # bits transmitted ≈ 2.2 Mbits

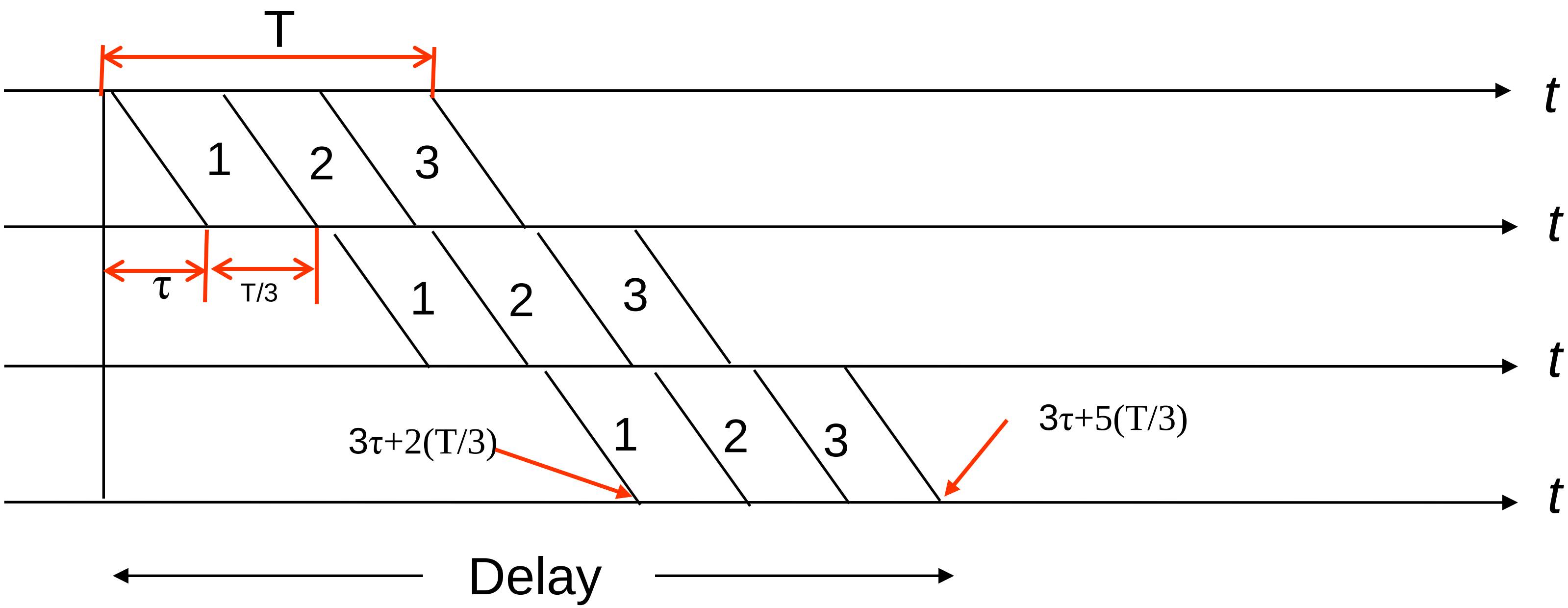
Connectionless Packet Switching - Datagram

- Messages broken into smaller units (packets)
- Source & destination addresses in packet header
- Connectionless, packets routed independently (datagram)
- Packet may arrive out of order
- Pipelining of packets across network can reduce delay, increase throughput
- Lower delay than message switching, suitable for interactive traffic



Packet Switching Delay

Assume three packets corresponding to one message traverse same path

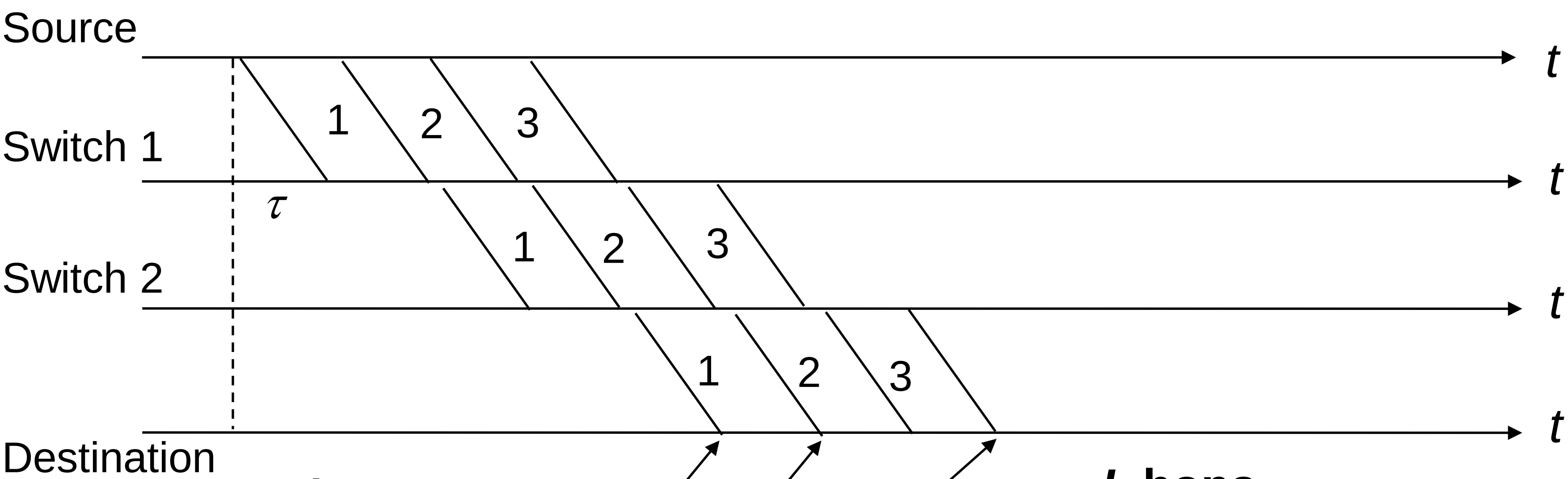


Minimum Delay = $3\tau + 5(T/3)$ (single path assumed)

Additional queueing delays possible at each link

Packet pipelining enables message to arrive sooner

Delay for k-Packet Message over L Hops



3 hops

$3\tau + 2(T/3)$ first bit received

$3\tau + 3(T/3)$ first bit released

$3\tau + 5(T/3)$ last bit released

L hops

$L\tau + (L-1)P$ first bit received

P: transmission delay / packet

$L\tau + LP$ first bit released

$L\tau + LP + (k-1)P$ last bit released

where $T = kP$

Routing Tables in Datagram Networks

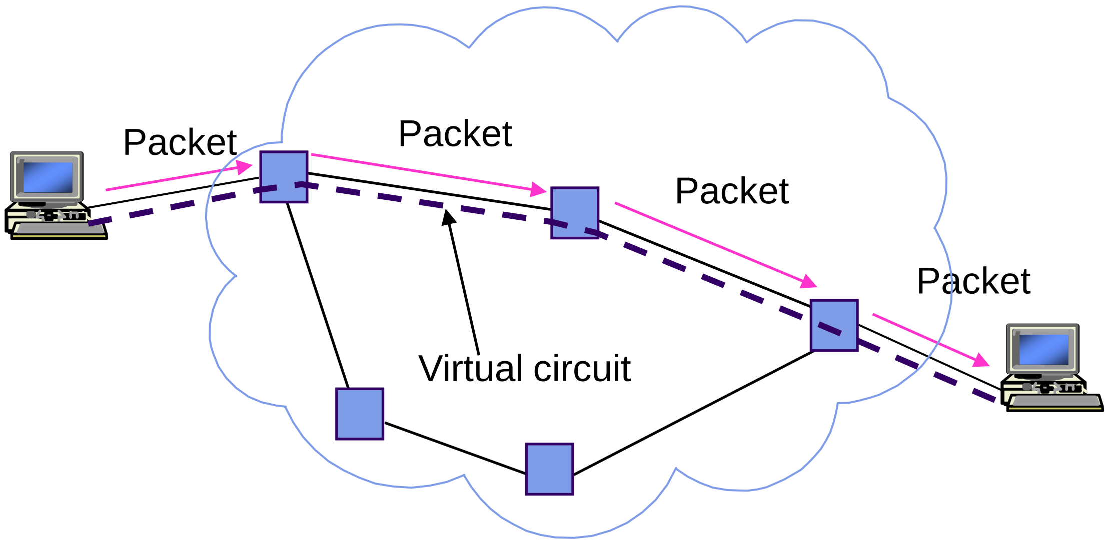
Destination address	Output port
0785	7
1345	12
1566	6
2458	12

- Route determined by table lookup
- Routing decision involves finding **next hop** in route to given destination
- Routing table has an entry for each destination specifying output port that leads to next hop
- Size of table becomes impractical for very large number of destinations

Example: Internet Routing

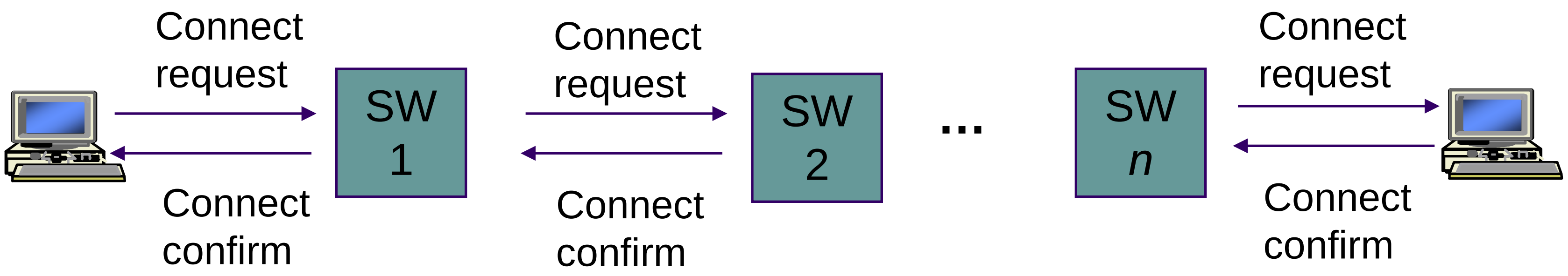
- Internet protocol uses datagram packet switching *across networks*
  - A packet arrives at a router…Router will do a table lookup of the packet destination address…if address is within its network packet will be forwarded to the appropriate output link…if address is not in the given network router will forward packet to a router of another network (next hop network) after performing suitable encapsulation …i.e. Networks are treated as data links
- Hosts have two-part IP address:
  - Network address + Host address
- Routers do table lookup on **network address**
  - This reduces size of routing table
- In addition, network addresses are assigned so that they can also be aggregated
  - To be discussed later

Packet Switching – Virtual Circuit



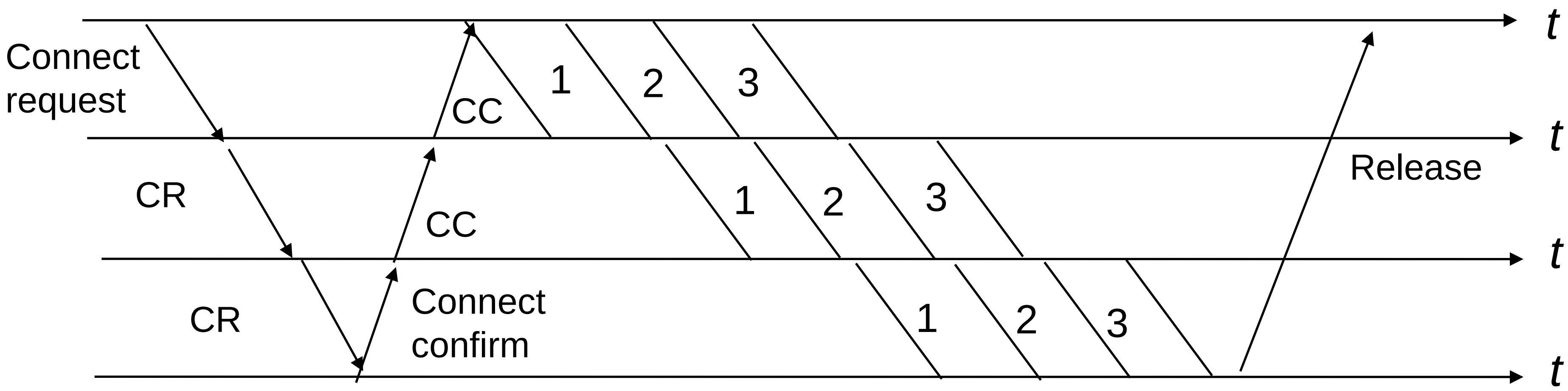
- Call set-up phase sets up pointers in fixed path along network
- All packets for a connection follow the same path
- Abbreviated header identifies connection on each link
- Packets queue for transmission
- Variable bit rates possible, negotiated during call set-up
- Delays variable, cannot be less than circuit switching

Connection Setup



- Signaling messages propagate as route is selected
- Signaling messages identify connection and set up tables in switches
- Typically a connection is identified by a local tag, Virtual Circuit Identifier (VCI)
- Each switch only needs to know how to relate an incoming tag in one input to an outgoing tag in the corresponding output
- Once tables are set, packets can flow along path

Connection Setup Delay



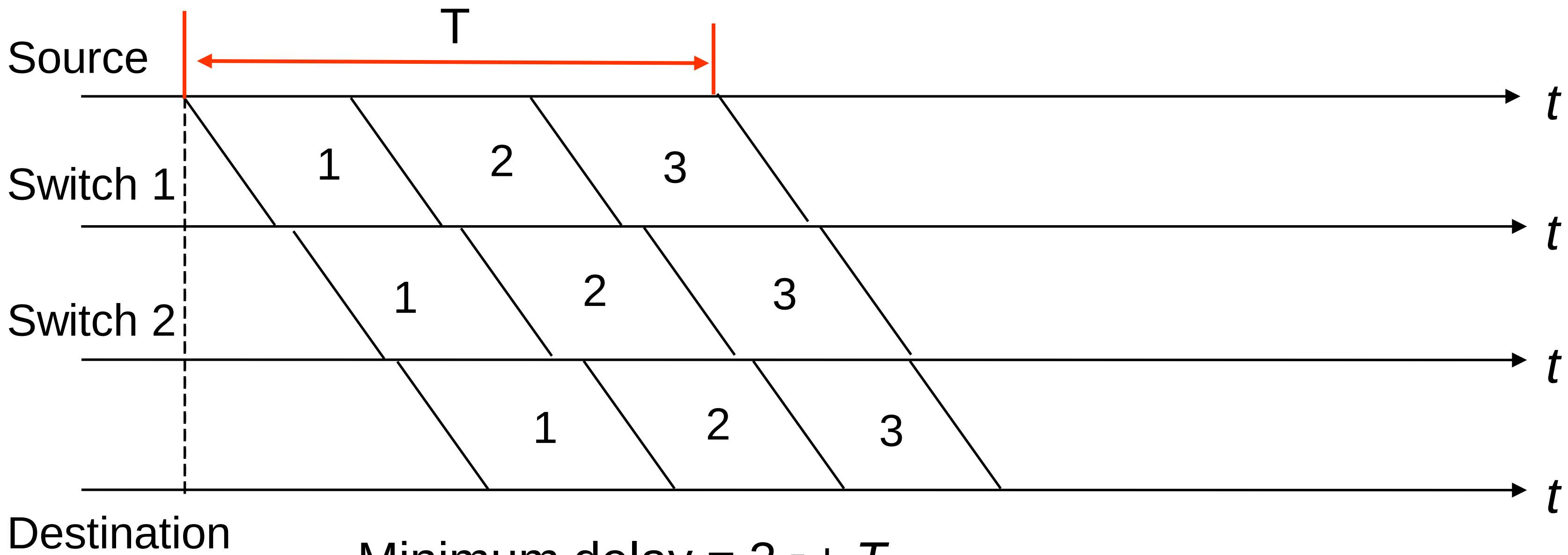
- Connection setup delay is incurred before any packet can be transferred
- Delay is acceptable for sustained transfer of large number of packets
- This delay may be unacceptably high if only a few packets are being transferred

Virtual Circuit Forwarding Tables

Input VCI	Output port	Output VCI
12	13	44
15	15	23
27	13	16
58	7	34

- Each input port of packet switch has a forwarding table
- Lookup entry for VCI of incoming packet
- Determine output port (next hop) and insert VCI for next link
- Very high speeds are possible
- Table can also include priority or other information about how packet should be treated

Cut-Through switching



Minimum delay = $3\tau + T$

- Some networks perform error checking on header only, so packet can be forwarded as soon as header is received & processed
- Delays reduced further with cut-through switching

Message vs. Packet Minimum Delay

- Message:

  $L\,\tau + L\,T \quad = L\,\tau + (L - 1)\,T + T$

- Packet

  $L\,\tau + L\,P + (k - 1)\,P \quad = L\,\tau + (L - 1)\,P + T$

- Cut-Through Packet (Immediate forwarding after header)

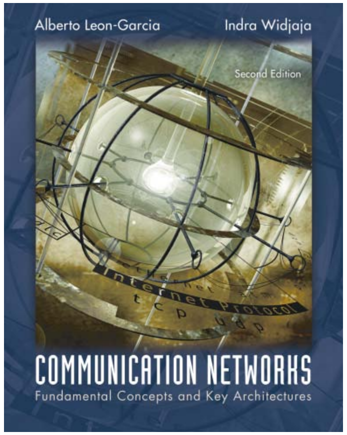
  $= L\,\tau + T$

  Above neglect header processing delays

Example: ATM Networks

- All information mapped into short fixed-length packets called *cells*
- Connections set up across network
  - Virtual circuits established across networks
  - Tables setup at ATM switches
- Several types of network services offered
  - Constant bit rate connections
  - Variable bit rate connections

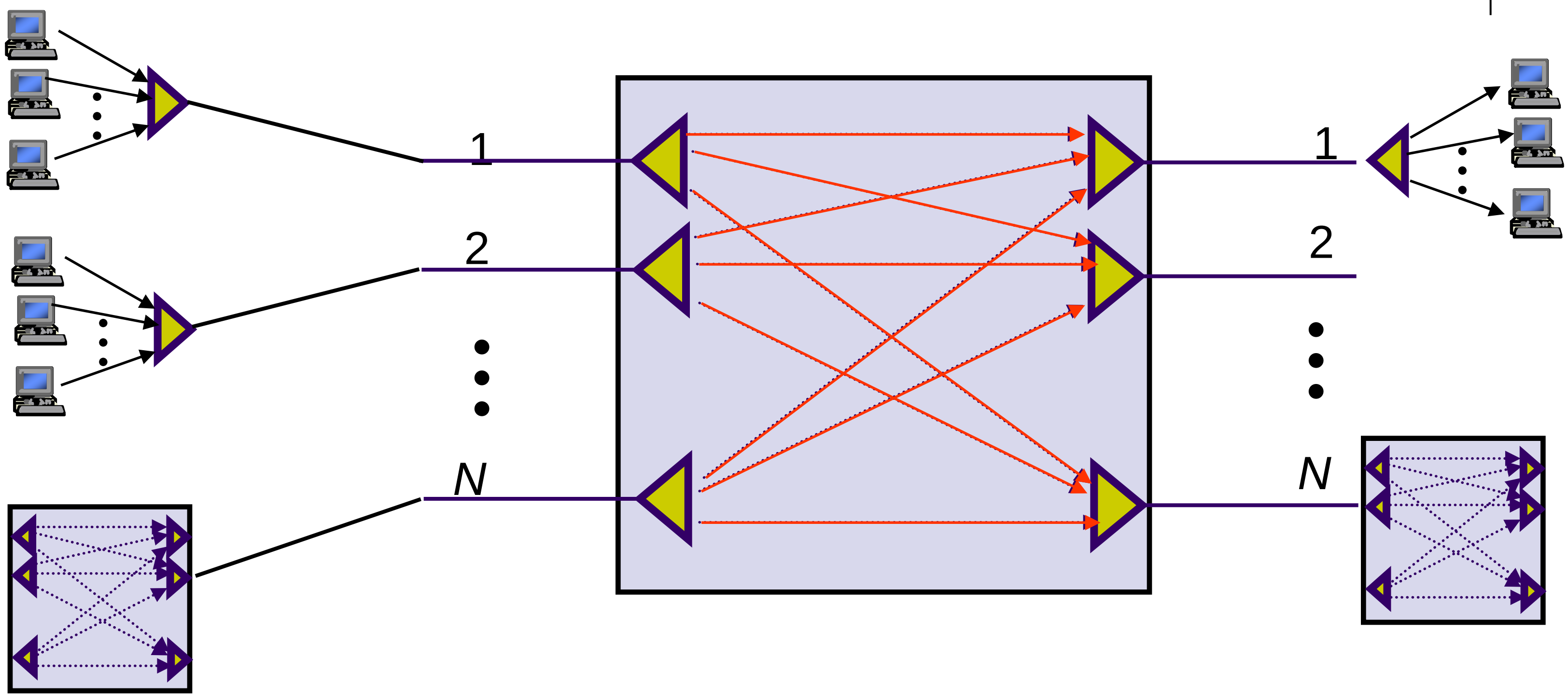
Chapter 7
Packet-Switching Networks



Datagrams and Virtual Circuits

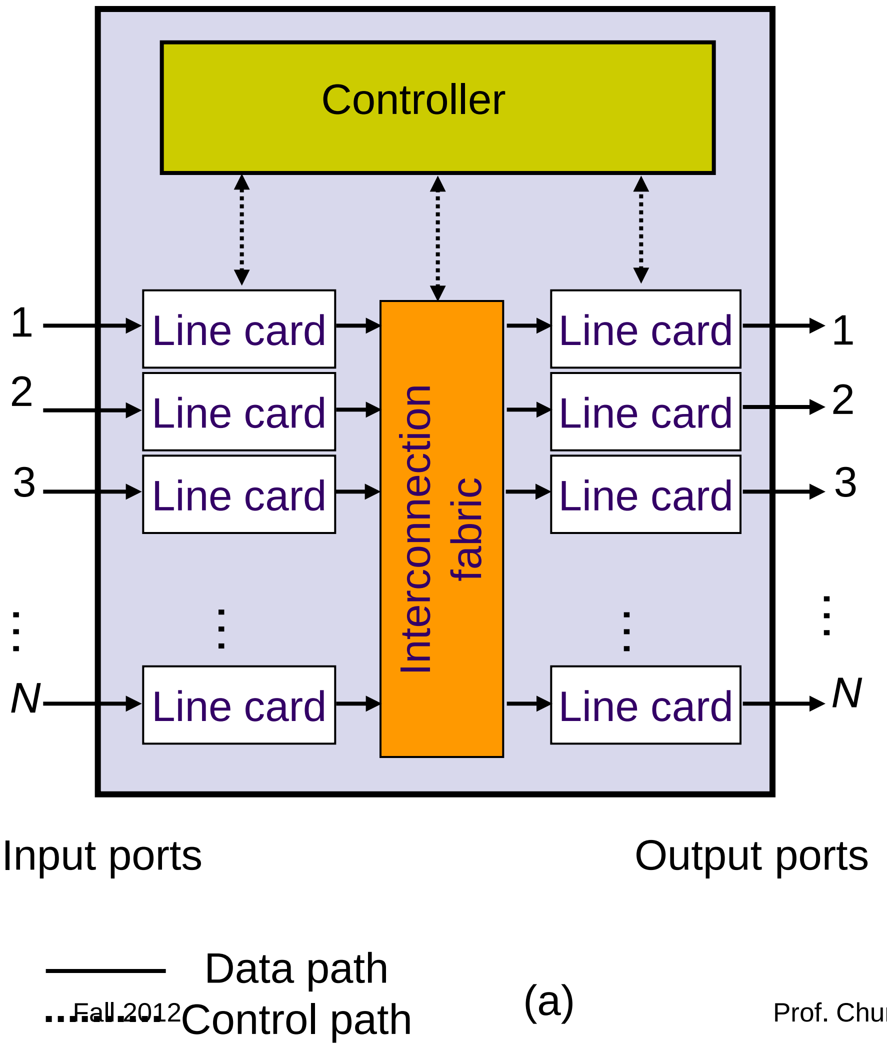
Structure of a Packet Switch

Packet Switch: Intersection where Traffic Flows Meet



- Inputs contain multiplexed flows from access muxs & other packet switches
- Flows demultiplexed at input, routed and/or forwarded to output ports
- Packets buffered, prioritized, and multiplexed on output lines

Generic Packet Switch

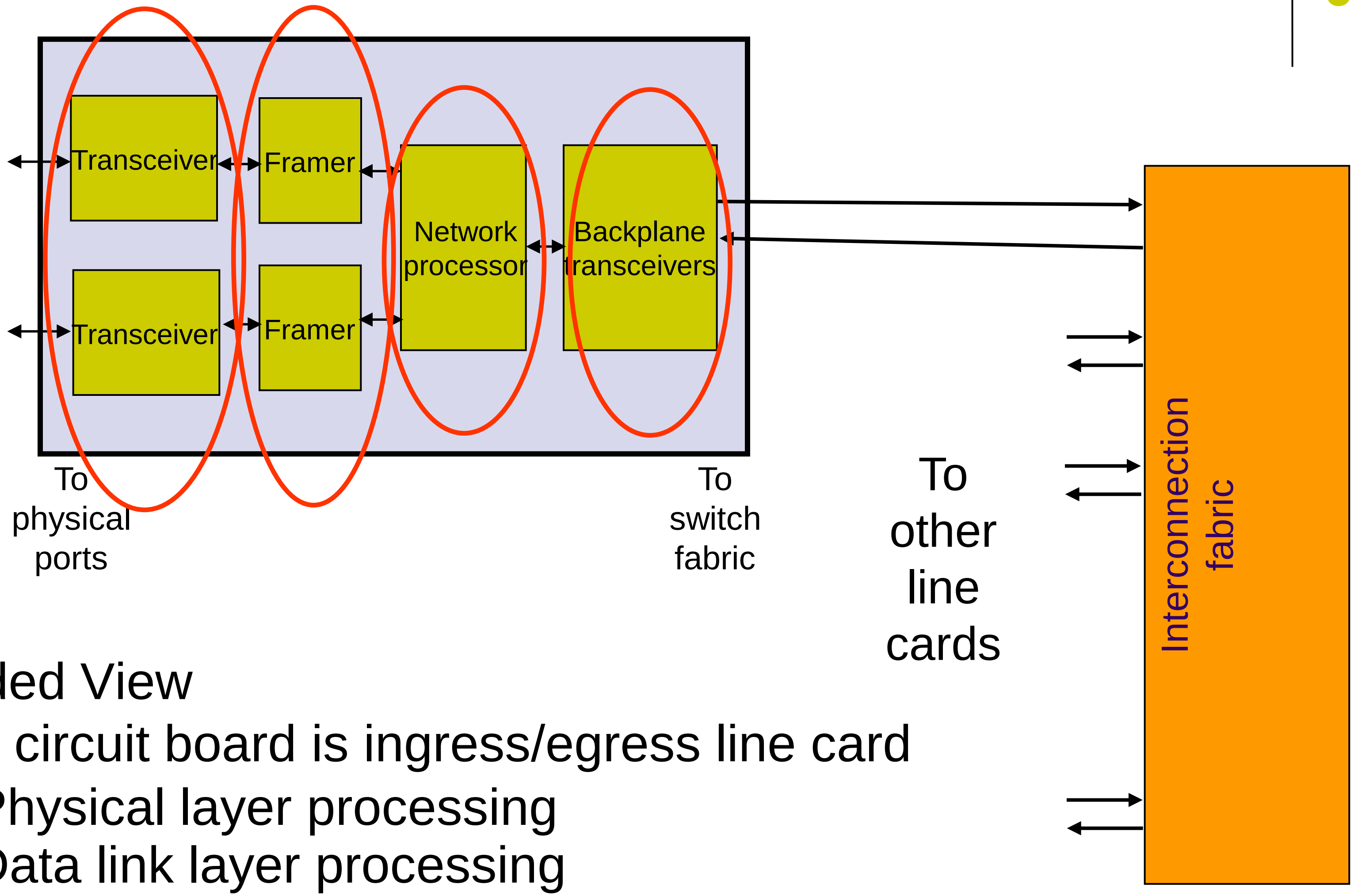


(a)

"Unfolded" View of Switch

- Ingress Line Cards
  - Header processing
  - Demultiplexing
  - Routing in large switches
- Controller
  - Routing protocols
  - Signalling & resource allocation
- Interconnection Fabric
  - Transfer packets between line cards
- Egress Line Cards
  - Scheduling & priority
  - Multiplexing

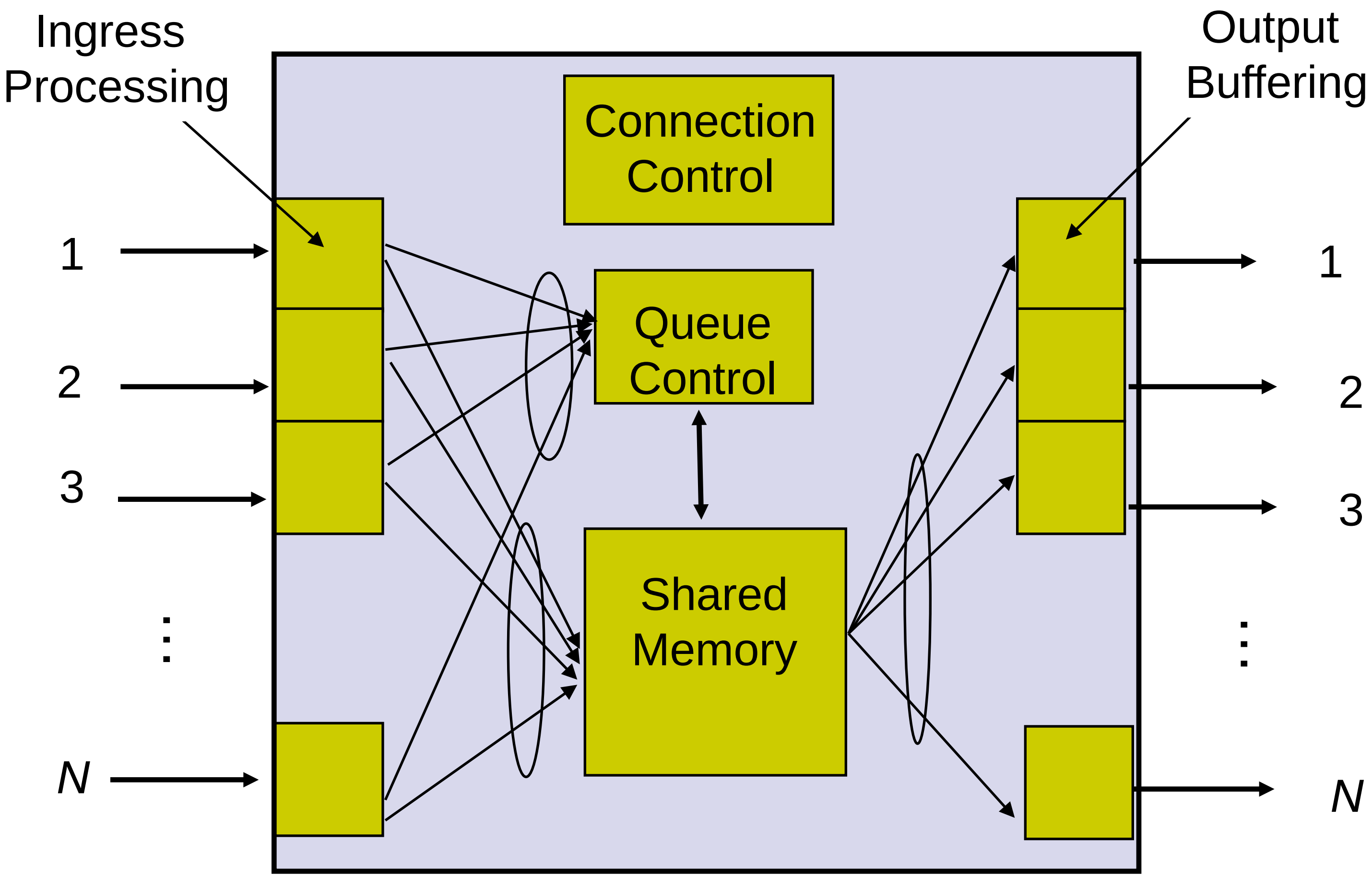
Line Cards



Folded View

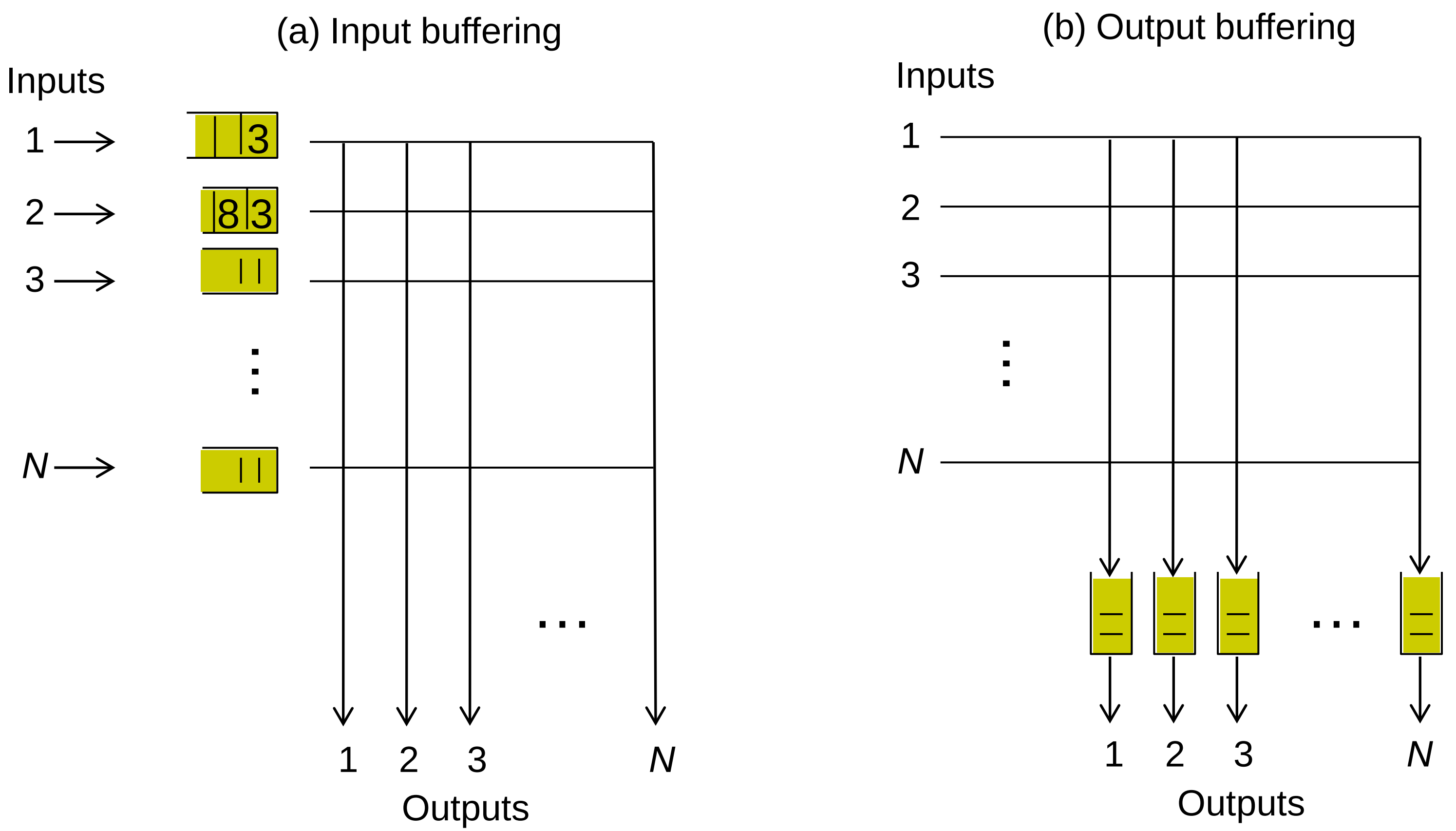
- 1 circuit board is ingress/egress line card
- Physical layer processing
- Data link layer processing
- Network header processing
- Physical layer across fabric + framing

Shared Memory Packet Switch



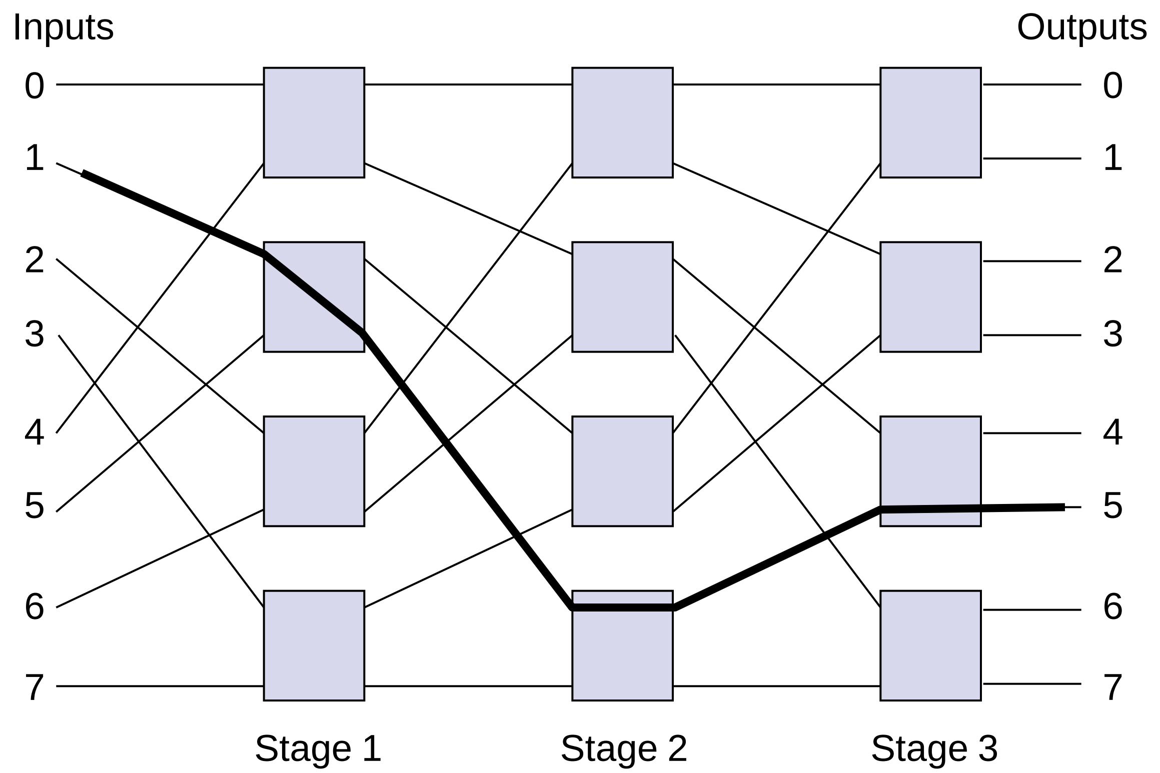
Small switches can be built by reading/writing into shared memory

Crossbar Switches



- Large switches built from crossbar & multistage space switches
- Requires centralized controller/scheduler (who sends to whom when)
- Can buffer at input, output, or both (performance vs complexity)

Self-Routing Switches



- Self-routing switches do not require controller
- Output port number determines route
- 101 → (1) lower port, (2) upper port, (3) lower port